

Transaction Costs, Trading Volume, and the Liquidity Premium*

Stefan Gerhold [†]

Paolo Guasoni [‡]

Johannes Muhle-Karbe [§]

Walter Schachermayer [¶]

February 20, 2012

Abstract

In a market with one safe and one risky asset, an investor with a long horizon, constant investment opportunities, and constant relative risk aversion trades with small proportional transaction costs. We derive explicit formulas for the optimal investment policy, its implied welfare, liquidity premium, and trading volume. At the first order, the liquidity premium equals the spread, times share turnover, times a universal constant. Results are robust to consumption and finite horizons. We exploit the equivalence of the transaction cost market to another frictionless market, with a shadow risky asset, in which investment opportunities are stochastic. The shadow price is also found explicitly.

Mathematics Subject Classification: (2010) 91G10, 91G80.

JEL Classification: G11, G12.

Keywords: transaction costs, long-run, portfolio choice, liquidity premium, trading volume.

1 Introduction

If risk aversion and investment opportunities are constant — and frictions are absent — investors should hold a constant mix of safe and risky assets (Markowitz, 1952; Merton, 1969, 1971). Transaction costs substantially change this statement, casting some doubt on its far-reaching implications.¹ Even the small spreads that are present in the most liquid markets entail wide oscillations in portfolio weights, which imply variable risk premia.

*For helpful comments, we thank Maxim Bichuch, George Constantinides, Aleš Černý, Mark Davis, Ioannis Karatzas, Ren Liu, Marcel Nutz, Scott Robertson, Johannes Ruf, Mihai Sirbu, Mete Soner, Gordan Zitković, and seminar participants at Ascona, MFO Oberwolfach, Columbia University, Princeton University, University of Oxford, CAU Kiel, London School of Economics, University of Michigan, TU Vienna, and the ICIAM meeting in Vancouver.

[†]Technische Universität Wien, Institut für Wirtschaftsmathematik, Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria, email sgerhold@fam.tuwien.ac.at. Partially supported by the Austrian Federal Financing Agency and the Christian-Doppler-Gesellschaft (CDG).

[‡]Boston University, Department of Mathematics and Statistics, 111 Cummington Street, Boston, MA 02215, USA, and Dublin City University, School of Mathematical Sciences, Glasnevin, Dublin 9, Ireland, email guasoni@bu.edu. Tel. +1 617 353 4992. Fax +1 617 353 8100. Partially supported by the ERC (278295), NSF (DMS-0807994, DMS-1109047), SFI (07/MI/008, 07/SK/M1189, 08/SRC/FMC1389), and FP7 (RG-248896).

[§]Corresponding author. ETH Zürich, Departement Mathematik, Sägmistrasse 101, CH-8092, Zürich, Switzerland, email johannes.muhle-karbe@math.ethz.ch. Partially supported by the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management), of the Swiss National Science Foundation (SNF).

[¶]Universität Wien, Fakultät für Mathematik, Nordbergstrasse 15, A-1090 Wien, Austria, email walter.schachermayer@univie.ac.at. Partially supported by the Austrian Science Fund (FWF) under grant P19456, the European Research Council (ERC) under grant FA506041, the Vienna Science and Technology Fund (WWTF) under grant MA09-003, and by the Christian-Doppler-Gesellschaft (CDG).

¹Constantinides (1986) finds that “transaction costs have a first-order effect on the assets’ demand.” Liu and Loewenstein (2002) note that “even small transaction costs lead to dramatic changes in the optimal behavior for an investor: from continuous trading to virtually buy-and-hold strategies.” Luttmer (1996) shows how small transaction costs help resolve asset pricing puzzles.

This paper studies a tractable benchmark of portfolio choice under transaction costs, with constant investment opportunities, summarized by a safe rate r , and a risky asset with volatility σ and expected excess return $\mu > 0$, which trades at a bid (selling) price $(1 - \varepsilon)S_t$ equal to a constant fraction $(1 - \varepsilon)$ of the ask (buying) price S_t . Our analysis is based on the model of Dumas and Luciano (1991), which concentrates on long-run asymptotics to gain in tractability. In their framework, we find explicit solutions for the optimal policy, welfare, liquidity premium² and trading volume, in terms of model parameters, and of an additional quantity, the *gap*, identified as the solution to a scalar equation. For all these quantities, we derive closed-form asymptotics, in terms of model parameters only, for small transaction costs.

We uncover novel relations among the liquidity premium, trading volume, and transaction costs. First, we show that share turnover (ShTu), the liquidity premium (LiPr), and the bid-ask spread ε satisfy the following asymptotic relation:

$$\text{LiPr} \approx \frac{3}{4}\varepsilon \text{ ShTu}. \quad (1.1)$$

This relation is universal, as it involves neither market nor preference parameters. Also, because it links the liquidity premium, which is unobservable, with spreads and share turnover, which are observable, this relation can help estimate the liquidity premium using data on trading volume.

Second, we find that the liquidity premium behaves very differently in the presence of leverage. In the no-leverage regime, the liquidity premium is an order of magnitude smaller than the spread (Constantinides, 1986), as unlevered investors respond to transaction costs by trading infrequently. With leverage, however, the liquidity premium increases quickly, because rebalancing a levered position entails high transaction costs, even under the optimal trading policy.

Third, we obtain the first continuous-time benchmark for trading volume, with explicit formulas for share and wealth turnover. Trading volume is an elusive quantity for frictionless models, in which turnover is typically infinite in any time interval.³ In the absence of leverage, our results imply low trading volume compared to the levels observed in the market. Of course, our model can only explain trading generated by portfolio rebalancing, and not by other motives such as market timing, hedging, and life-cycle investing.

Moreover, welfare, the liquidity premium, and trading volume depend on the market parameters (μ, σ) only through the mean-variance ratio μ/σ^2 if measured in *business time*, that is, using a clock that ticks at the speed of the market's variance σ^2 . Thus, in usual calendar time, all these quantities are functions of μ/σ^2 , multiplied by the variance σ^2 .

Our main implication for portfolio choice is that a symmetric, stationary policy is optimal for a long horizon, and it is robust, at the first order, both to intermediate consumption, and to a finite horizon. Indeed, we show that the no-trade region is perfectly symmetric with respect to the Merton proportion $\pi_* = \mu/\gamma\sigma^2$, if trading boundaries are expressed with trading prices, that is, if the buy boundary π_- is computed from the ask price, and the sell boundary π_+ from the bid price.

Since in a frictionless market the optimal policy is independent both of intermediate consumption and of the horizon (Merton, 1971), our results entail that these two features are robust to small frictions. However plausible these conclusions may seem, the literature so far has offered

² That is, the amount of excess return the investor is ready to forgo to trade the risky asset without transaction costs.

³The empirical literature has long been aware of this theoretical vacuum: Gallant, Rossi and Tauchen (1992) reckon that “The intrinsic difficulties of specifying plausible, rigorous, and implementable models of volume and prices are the reasons for the informal modeling approaches commonly used.” Lo and Wang (2000) note that “although most models of asset markets have focused on the behavior of returns [...] their implications for trading volume have received far less attention.”

diverse views on these issues (cf. Davis and Norman (1990); Dumas and Luciano (1991); Liu and Loewenstein (2002)). More importantly, robustness to the horizon implies that the long-horizon approximation, made for the sake of tractability, is reasonable and relevant. For typical parameter values, we see that our optimal strategy is nearly optimal already for horizons as short as two years.

A key idea for our results — and for their proof — is the equivalence between a market with transaction costs and constant investment opportunities, and another *shadow* market, without transaction costs, but with stochastic investment opportunities driven by a state variable. This state variable is the ratio between the investor's risky and safe weights, which tracks the location of the portfolio within the trading boundaries, and affects both the volatility and the expected return of the shadow risky asset.

The paper is organized as follows: Section 2 introduces the portfolio choice problem and states the main results. The model's main implications are discussed in Section 3, and the main results are derived heuristically in Section 4. Section 5 concludes, and all proofs are in the appendix.

2 Model and Main Result

Consider a market with a safe asset earning an interest rate r , i.e. $S_t^0 = e^{rt}$, and a risky asset, trading at ask (buying) price S_t following geometric Brownian motion,

$$dS_t/S_t = (\mu + r)dt + \sigma dW_t.$$

Here, W_t is a standard Brownian motion, $\mu > 0$ is the expected excess return,⁴ and $\sigma > 0$ is the volatility. The corresponding bid (selling) price is $(1 - \varepsilon)S_t$, where $\varepsilon \in (0, 1)$ represents the relative bid-ask spread.

A self-financing *trading strategy* is a two-dimensional, predictable process (φ_t^0, φ_t) of finite variation, such that φ_t^0 and φ_t represent the number of units in the safe and risky asset at time t , and the initial number of units is $(\varphi_{0-}^0, \varphi_{0-}) \in \mathbb{R}_+^2 \setminus \{0, 0\}$. Writing $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ as the difference between the cumulative number of shares bought (φ_t^\uparrow) and sold (φ_t^\downarrow) by time t , the *self-financing condition* relates the dynamics of φ^0 and φ via

$$d\varphi_t^0 = -\frac{S_t}{S_t^0}d\varphi_t^\uparrow + (1 - \varepsilon)\frac{S_t}{S_t^0}d\varphi_t^\downarrow. \quad (2.1)$$

As in Dumas and Luciano (1991), the investor maximizes the equivalent safe rate of power utility, an optimization objective that also proved useful with constraints on leverage (Grossman and Vila, 1992) and drawdowns (Grossman and Zhou, 1993).

Definition 2.1. A trading strategy (φ_t^0, φ_t) is admissible if its liquidation value is positive, in that:

$$\Xi_t^\varphi = \varphi_t^0 S_t^0 + (1 - \varepsilon)S_t \varphi_t^\uparrow - \varphi_t^\downarrow S_t \geq 0, \quad \text{a.s. for all } t \geq 0.$$

An admissible strategy (φ_t^0, φ_t) is long-run optimal if it maximizes the equivalent safe rate

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \log E [(\Xi_T^\varphi)^{1-\gamma}]^{\frac{1}{1-\gamma}} \quad (2.2)$$

over all admissible strategies, where $1 \neq \gamma > 0$ denotes the investor's relative risk aversion.⁵

⁴A negative excess return leads to a similar treatment, but entails buying as prices rise, rather than fall. For the sake of clarity, the rest of the paper concentrates on the more relevant case of a positive μ .

⁵The limiting case $\gamma \rightarrow 1$ corresponds to logarithmic utility, studied by Taksar, Klass and Assaf (1988) and Gerhold, Muhle-Karbe and Schachermayer (2011b). Theorem 2.2 remains valid for logarithmic utility setting $\gamma = 1$.

Our main result is the following:

Theorem 2.2. *An investor with constant relative risk aversion $\gamma > 0$ trades to maximize (2.2). Then, for small transaction costs $\varepsilon > 0$:*

i) (Equivalent Safe Rate)

For the investor, trading the risky asset with transaction costs is equivalent to leaving all wealth in a hypothetical safe asset, which pays the higher equivalent safe rate:

$$\text{ESR} = r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2}, \quad (2.3)$$

where the gap λ is defined in iv) below.

ii) (Liquidity Premium)

Trading the risky asset with transaction costs is equivalent to trading a hypothetical asset, at no transaction costs, with the same volatility σ , but with lower expected excess return $\sqrt{\mu^2 - \lambda^2}$. Thus, the liquidity premium is

$$\text{LiPr} = \mu - \sqrt{\mu^2 - \lambda^2}. \quad (2.4)$$

iii) (Trading Policy)

It is optimal to keep the fraction of wealth in the risky asset within the buy and sell boundaries

$$\pi_- = \frac{\mu - \lambda}{\gamma\sigma^2}, \quad \pi_+ = \frac{\mu + \lambda}{\gamma\sigma^2}, \quad (2.5)$$

where π_- and π_+ are computed with ask and bid prices, respectively.⁶

iv) (Gap)

λ is the unique value for which the solution of the initial value problem

$$\begin{aligned} w'(x) + (1 - \gamma)w(x)^2 + \left(\frac{2\mu}{\sigma^2} - 1\right)w(x) - \gamma\left(\frac{\mu - \lambda}{\gamma\sigma^2}\right)\left(\frac{\mu + \lambda}{\gamma\sigma^2}\right) &= 0 \\ w(0) &= \frac{\mu - \lambda}{\gamma\sigma^2}, \end{aligned}$$

also satisfies the terminal value condition:

$$w\left(\log\left(\frac{u(\lambda)}{l(\lambda)}\right)\right) = \frac{\mu + \lambda}{\gamma\sigma^2}, \quad \text{where} \quad \frac{u(\lambda)}{l(\lambda)} = \frac{1}{(1 - \varepsilon)} \frac{(\mu + \lambda)(\mu - \lambda - \gamma\sigma^2)}{(\mu - \lambda)(\mu + \lambda - \gamma\sigma^2)}.$$

In view of the explicit formula for $w(x, \lambda)$ in Lemma B.1 below, this is a scalar equation for λ .

v) (Trading Volume)

Share turnover, defined as shares traded $d\|\varphi\|_t = d\varphi_t^\uparrow + d\varphi_t^\downarrow$ divided by shares held $|\varphi_t|$, has the long-term average

$$\text{ShTu} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{d\|\varphi\|_t}{|\varphi_t|} = \frac{\sigma^2}{2} \left(\frac{2\mu}{\sigma^2} - 1\right) \left(\frac{1 - \pi_-}{(u(\lambda)/l(\lambda))^{\frac{2\mu}{\sigma^2} - 1} - 1} - \frac{1 - \pi_+}{(u(\lambda)/l(\lambda))^{1 - \frac{2\mu}{\sigma^2}} - 1} \right).$$

⁶This optimal policy is not necessarily unique, in that its long-run performance is also attained by trading arbitrarily for a finite time, and then switching to the above policy. However, in related frictionless models, as the horizon increases, the optimal (finite-horizon) policy converges to a stationary policy, such as the one considered here (see, e.g., Dybvig, Rogers and Back (1999)). Dai and Yi (2009) obtain similar results in a model with proportional transaction costs, formally passing to a stationary version of their control problem PDE.

Wealth turnover, defined as wealth traded divided by wealth held, has long term-average:⁷

$$\begin{aligned}\text{WeTu} &= \lim_{T \rightarrow \infty} \frac{1}{T} \left(\int_0^T \frac{(1-\varepsilon)S_t d\varphi_t^\downarrow}{\varphi_t^0 S_t^0 + \varphi_t(1-\varepsilon)S_t} + \int_0^T \frac{S_t d\varphi_t^\uparrow}{\varphi_t^0 S_t^0 + \varphi_t S_t} \right) \\ &= \frac{\sigma^2}{2} \left(\frac{2\mu}{\sigma^2} - 1 \right) \left(\frac{\pi_- (1 - \pi_-)}{(u(\lambda)/l(\lambda))^{\frac{2\mu}{\sigma^2}-1} - 1} - \frac{\pi_+ (1 - \pi_+)}{(u(\lambda)/l(\lambda))^{1-\frac{2\mu}{\sigma^2}} - 1} \right).\end{aligned}$$

vi) (Asymptotics)

Setting $\pi_* = \mu/\gamma\sigma^2$, the following expansions in terms of the bid-ask spread ε hold:⁸

$$\lambda = \gamma\sigma^2 \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{1/3} \varepsilon^{1/3} + O(\varepsilon). \quad (2.6)$$

$$\text{ESR} = r + \frac{\mu^2}{2\gamma\sigma^2} - \frac{\gamma\sigma^2}{2} \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{2/3} \varepsilon^{2/3} + O(\varepsilon^{4/3}). \quad (2.7)$$

$$\text{LiPr} = \frac{\mu}{2\pi_*^2} \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{2/3} \varepsilon^{2/3} + O(\varepsilon^{4/3}). \quad (2.8)$$

$$\pi_\pm = \pi_* \pm \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{1/3} \varepsilon^{1/3} + O(\varepsilon). \quad (2.9)$$

$$\text{ShTu} = \frac{\sigma^2}{2} (1 - \pi_*)^2 \pi_* \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{-1/3} \varepsilon^{-1/3} + O(\varepsilon^{1/3}). \quad (2.10)$$

$$\text{WeTu} = \frac{2\gamma\sigma^2}{3} \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{2/3} \varepsilon^{-1/3} + O(\varepsilon^{1/3}). \quad (2.11)$$

In summary, our optimal trading policy, and its resulting welfare, liquidity premium, and trading volume are all simple functions of investment opportunities (r, μ, σ) , preferences (γ) , and the gap λ . The gap does not admit an explicit formula in terms of the transaction cost parameter ε , but is determined through the implicit relation in *iv*), and has the asymptotic expansion in *vi*), from which all other asymptotic expansions follow through the explicit formulas.

The frictionless markets with constant investment opportunities in items *i*) and *ii*) of Theorem 2.2 are equivalent to the market with transaction costs in terms of equivalent safe rates. Nevertheless, the corresponding optimal policies are very different, requiring no or incessant re-balancing, respectively, in the frictionless markets but only finite trading volume with transaction costs.

By contrast, the shadow price, which is key in the derivation of our results, is a fictitious risky asset, with price evolving within the bid-ask spread, that is equivalent to the transaction cost market in terms of both welfare *and* the optimal policy:

Theorem 2.3. *The policy in Theorem 2.2 iii) and the equivalent safe rate in Theorem 2.2 i) are also optimal for a frictionless asset with shadow price \tilde{S} , which always lies within the bid-ask spread, and coincides with the trading price at times of trading for the optimal policy. The shadow price satisfies*

$$d\tilde{S}_t/\tilde{S}_t = (\tilde{\mu}(Y_t) + r)dt + \tilde{\sigma}(Y_t)dW_t, \quad (2.12)$$

⁷The number of shares is written as the difference $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ of the cumulative shares bought (resp. sold), and wealth is evaluated at trading prices, i.e., at the bid price $(1-\varepsilon)S_t$ when selling, and at the ask price S_t when buying.

⁸Algorithmic calculations can deliver terms of arbitrarily high order.

for the deterministic functions $\tilde{\mu}(\cdot)$ and $\tilde{\sigma}(\cdot)$ given explicitly in Lemma C.1. The state variable $Y_t = \log(\varphi_t S_t / (l\varphi_t^0 S_t^0))$ represents the logarithm of the ratio of risky and safe positions, which follows a Brownian motion with drift, reflected to remain in the interval $[0, \log(u(\lambda)/l(\lambda))]$, i.e.,

$$dY_t = (\mu - \sigma^2/2)dt + \sigma dW_t + dL_t - dU_t. \quad (2.13)$$

Here, L_t and U_t are increasing processes, proportional to the cumulative purchases and sales, respectively (cf. (C.10) below). In the interior of the no-trade region, that is, when Y_t lies in $(0, \log(u(\lambda)/l(\lambda)))$, the numbers of units of the safe and risky asset are constant, and the state variable Y_t follows Brownian motion with drift. As Y_t reaches the boundary of the no-trade region, buying or selling takes place as to keep it within $[0, \log(u(\lambda)/l(\lambda))]$.

In view of Theorem 2.3, trading with constant investment opportunities and proportional transaction costs is equivalent to trading in a fictitious frictionless market with stochastic investment opportunities, which vary with the location of the investor's portfolio in the no-trade region.

3 Implications

3.1 Trading Strategies

Equation (2.5) implies that trading boundaries are symmetric around the frictionless Merton proportion $\pi_* = \mu/\gamma\sigma^2$. At first glance, this seems to contradict previous studies (e.g., Liu and Loewenstein (2002), Shreve and Soner (1994)), which emphasize how these boundaries are asymmetric, and may even fail to include the Merton proportion. These papers employ a common reference price (the average of the bid and ask prices) to evaluate both boundaries. By contrast, we express trading boundaries using trading prices (i.e., the ask price for the buy boundary, and the bid price for the sell boundary). This simple convention unveils the natural symmetry of the optimal policy, and explains asymmetries as figments of notation – even in their models. To see this, denote by π'_- and π'_+ the buy and sell boundaries in terms of the ask price. These papers prove the bounds (Shreve and Soner (1994, equations (11.4) and (11.6)) in a infinite-horizon model with consumption and Liu and Loewenstein (2002, equations (22), (23)) in a finite-horizon model)

$$\pi'_- < \frac{\mu}{\gamma\sigma^2} \quad \text{and} \quad \frac{\mu}{\gamma\sigma^2(1-\varepsilon) + \varepsilon\mu} < \pi'_+ < \frac{\mu}{\frac{1}{2}\gamma\sigma^2(1-\varepsilon) + \varepsilon\mu}. \quad (3.1)$$

With trading prices (i.e., substituting $\pi_- = \pi'_-$ and $\pi_+ = \frac{1-\varepsilon}{1-\varepsilon\pi'_+}\pi'_+$) these bounds become

$$\pi_- < \frac{\mu}{\gamma\sigma^2} < \pi_+ < 2\frac{\mu}{\gamma\sigma^2}, \quad (3.2)$$

whence the Merton proportion always lies between π_- and π_+ .

Because our model excludes intermediate consumption, we compare our trading boundaries with those obtained by Davis and Norman (1990) and Shreve and Soner (1994) in the consumption model of Magill and Constantinides (1976). The asymptotic expansions of Janeček and Shreve (2004) make this comparison straightforward.

With or without consumption, the trading boundaries coincide at the first-order. This fact has a clear economic interpretation: the separation between consumption and investment, which holds in a frictionless model with constant investment opportunities, is a robust feature of frictionless models, because it still holds, *at the first order*, even with transaction costs. Put differently, if investment opportunities are constant, consumption has only a second order effect for investment decisions, in spite of the large no-trade region implied by transaction costs. Figure 1 shows that our bounds are very close to those obtained in the model of Davis and Norman (1990) for bid-ask spreads below 1%, but start diverging for larger values.

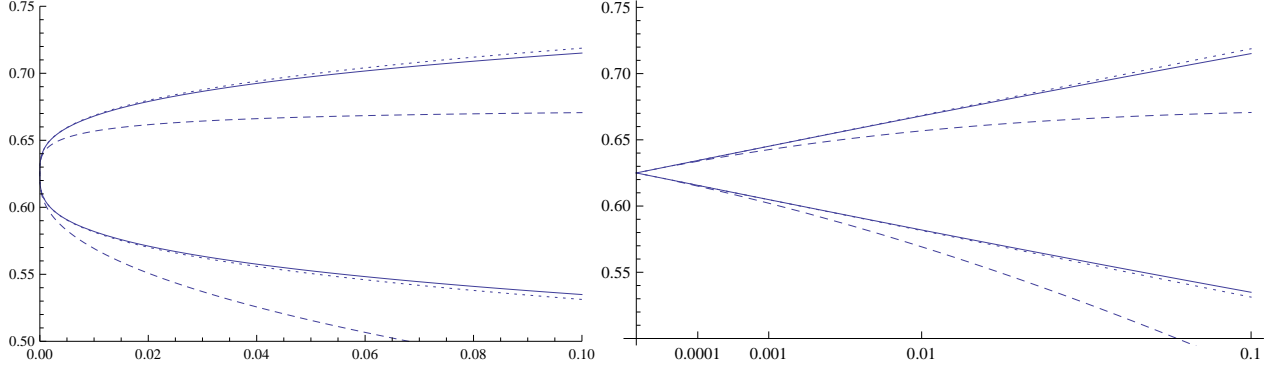


Figure 1: Buy (lower) and sell (upper) boundaries (vertical axis, as risky weights) as functions of the spread ε , in linear scale (left panel) and cubic scale (right panel). The plot compares the approximate weights from the first term of the expansion (dotted), the exact optimal weights (solid), and the boundaries found by Davis and Norman (1990) in the presence of consumption (dashed). Parameters are $\mu = 8\%$, $\sigma = 16\%$, $\gamma = 5$, and a zero discount rate for consumption (for the dashed curve).

3.2 Business time and mean-variance ratio

In a frictionless market, the equivalent safe rate and the optimal policy are:

$$\text{ESR} = r + \frac{1}{2\gamma} \left(\frac{\mu}{\sigma} \right)^2 \quad \text{and} \quad \pi_* = \frac{\mu}{\gamma\sigma^2}.$$

This rate depends only on the safe rate r and the Sharpe ratio μ/σ . Investors are indifferent between two markets with identical safe rates and Sharpe ratios, because both markets lead to the same set of payoffs, even though a payoff is generated by different portfolios in the two markets. By contrast, the optimal portfolio depends only on the mean-variance ratio μ/σ^2 .

With transaction costs, Equation (2.6) shows that the asymptotic expansion of the gap per unit of variance λ/σ^2 only depends on the mean-variance ratio μ/σ^2 . Put differently, holding the mean-variance ratio μ/σ^2 constant, the expansion of λ is linear in σ^2 . In fact, not only the expansion but also the exact quantity has this property, since λ/σ^2 in *iv*) only depends on μ/σ^2 .

Consequently, the optimal policy in *iii*) only depends on the mean-variance ratio μ/σ^2 , as in the frictionless case. The equivalent safe rate, however, no longer solely depends on the Sharpe ratio μ/σ : investors are not indifferent between two markets with the same Sharpe ratio, because one market is more attractive than the other if it entails lower trading costs. As an extreme case, in one market it may be optimal to leave all wealth in the risky asset, eliminating any need to trade. Instead, the formulas in *i*), *ii*), and *v*) show that, like the gap per variance λ/σ^2 , the equivalent safe rate, the liquidity premium, and both share and wealth turnover only depend on μ/σ^2 , when measured per unit of variance. The interpretation is that these quantities are proportional to business time $\sigma^2 t$ (Ané and Geman, 2000), and the factor of σ^2 arises from measuring them in calendar time.

In the frictionless limit, the linearity in σ^2 and the dependence on μ/σ^2 cancel, and the result depends on the Sharpe ratio alone. For example, the equivalent safe rate becomes⁹

$$\frac{\sigma^2}{2} \left(\frac{\mu}{\sigma^2} \right)^2 = \frac{1}{2} \left(\frac{\mu}{\sigma} \right)^2.$$

⁹The other quantities are trivial: the gap and the liquidity premium become zero, while share and wealth turnover explode to infinity.

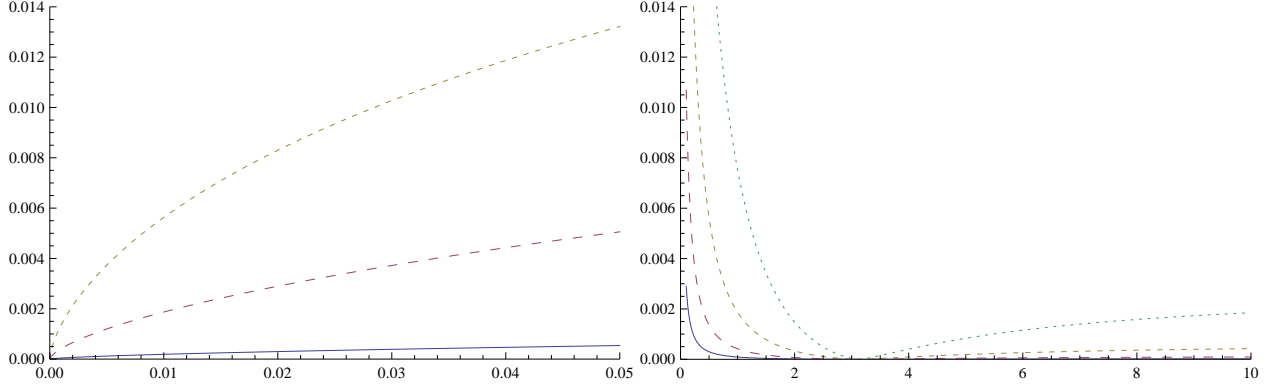


Figure 2: Left panel: liquidity premium (vertical axis) against the spread ε , for risk aversion γ equal to 5 (solid), 1 (long dashed), and 0.5 (short dashed). Right panel: liquidity premium (vertical axis) against risk aversion γ , for spread $\varepsilon = 0.01\%$ (solid), 0.1% (long dashed), 1% (short dashed), and 10% (dotted). Parameters are $\mu = 8\%$ and $\sigma = 16\%$.

3.3 Liquidity Premium

The liquidity premium (Constantinides, 1986) is the amount of expected excess return the investor is ready to forgo to trade the risky asset without transaction costs, as to achieve the same equivalent safe rate. Figure 2 plots the liquidity premium against the spread ε (left panel) and risk aversion γ (right panel).

The liquidity premium is exactly zero when the Merton proportion π_* is either zero or one. In these two limit cases, it is optimal not to trade at all, hence no compensation is required for the costs of trading. The liquidity premium is relatively low in the regime of no leverage ($0 < \pi_* < 1$), corresponding to $\gamma > \mu/\sigma^2$, confirming the results of Constantinides (1986), who reports liquidity premia one order of magnitude smaller than trading costs.

The leverage regime ($\gamma < \mu/\sigma^2$), however, shows a very different picture. As risk aversion decreases below the full-investment level $\gamma = \mu/\sigma^2$, the liquidity premium increases rapidly to infinity, as lower levels of risk aversion prescribe increasingly high leverage. The costs of rebalancing a levered position are high, and so are the corresponding liquidity premia.

The liquidity premium increases in spite of the increasing width of the no-trade region for larger leverage ratios. In other words, even as a less risk averse investor tolerates wider oscillations in the risky weight, this increased flexibility is not enough to compensate for the higher costs required to rebalance a more volatile portfolio.

3.4 Trading Volume

In the empirical literature (cf. Lo and Wang (2000) and the references therein), the most common measure of trading volume is share turnover, defined as number of shares traded divided by shares held or, equivalently, as the value of shares traded divided by value of shares held. In our model, turnover is positive only at the trading boundaries, while it is null inside the no-trade region. Since turnover, on average, grows linearly over time, we consider the long-term average of share turnover per unit of time, plotted in Figure 3 against risk aversion. Turnover is null at the full-investment level $\gamma = \mu/\sigma^2$, as no trading takes place in this case. Lower levels of risk aversion generate leverage, and trading volume increases rapidly, like the liquidity premium.

Share turnover does not decrease to zero as the risky weight decreases to zero for increasing risk

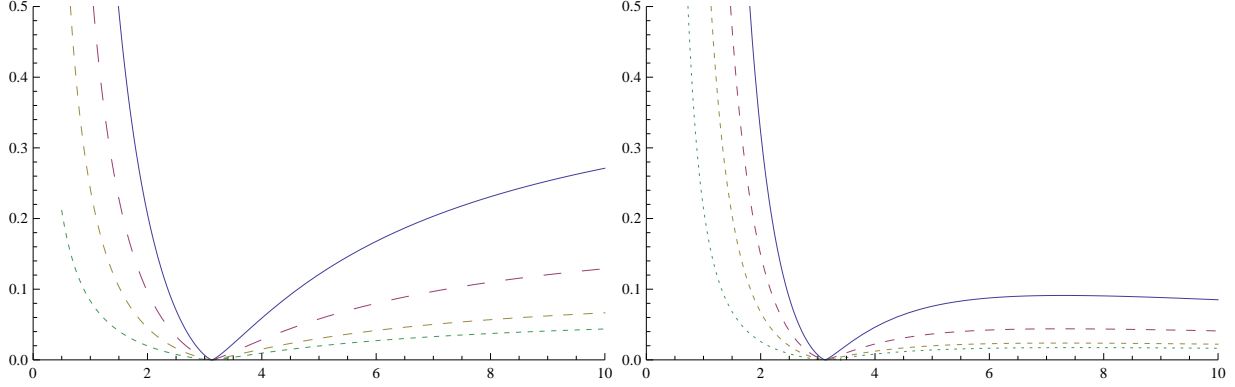


Figure 3: Trading volume (vertical axis, annual fractions traded), as share turnover (left panel) and wealth turnover (right panel), against risk aversion (horizontal axis), for spread $\varepsilon = 0.01\%$ (solid), 0.1% (long dashed), 1% (short dashed), and 10% (dotted). Parameters are $\mu = 8\%$ and $\sigma = 16\%$.

aversion γ . On the contrary, the first term in the asymptotic formula converges to a finite level. This phenomenon arises because more risk averse investors hold less risky assets (reducing volume), but also rebalance more frequently (increasing volume). As risk aversion increases, neither of these effects prevails, and turnover converges to a finite limit.

To better understand these properties, consider wealth turnover, defined as the value of shares traded, divided by total wealth (not by the value of shares held).¹⁰ Share and wealth turnover are qualitatively similar for low risk aversion, as the risky weight of wealth is larger, but they diverge as risk aversion increases and the risky weight declines to zero. Then, wealth turnover decreases to zero, whereas share turnover does not.

The levels of trading volume observed empirically imply very low values of risk aversion in our model. For example, Lo and Wang (2000) report in the NYSE-AMEX an average *weekly* turnover of 0.78% between 1962-1996, which corresponds to an approximate annual turnover above 40% . As Figure 3 shows, such a high level of turnover requires a risk aversion below 2, even for a very small spread of $\varepsilon = 0.01\%$. This phenomenon intensifies in the last two decades. As shown by Figure 4 turnover increases substantially from 1993 to 2010, with monthly averages of 20% typical from 2007 on, corresponding to an annual turnover of over 240% .

The overall implication is that portfolio rebalancing can generate substantial trading volume, but the model explains the trading volume observed empirically only with low risk aversion and high leverage. In a numerical study with risk aversion of six and spreads of 2% , Lynch and Tan (2011) also find that the resulting trading volume is too low, even allowing for labor income and predictable returns. Our analytical results are consistent with their findings, but indicate that substantially higher volume can be explained with lower risk aversion.

3.5 Volume, Spreads and the Liquidity Premium

The analogies between the comparative statics of the liquidity premium and trading volume suggest a close connection between these quantities. An inspection of the asymptotic formulas unveils the

¹⁰Technically, wealth is valued at the ask price at the buying boundary, and at the bid price at the selling boundary.

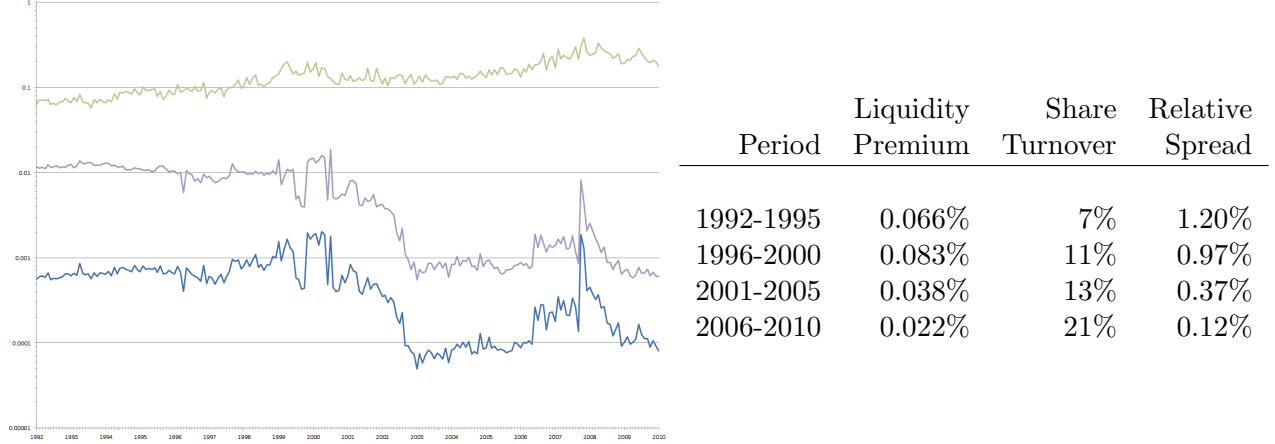


Figure 4: Left panel: share turnover (top), spread (center), and implied liquidity premium (bottom) in logarithmic scale, from 1992 to 2010. Right panel: monthly averages for share turnover, spread, and implied liquidity premium over subperiods. Spread and turnover are capitalization-weighted averages across securities in the monthly CRSP database with share codes 10, 11 that have nonzero bid, ask, volume and shares outstanding.

following relations:

$$\text{LiPr} = \frac{3}{4}\varepsilon\text{ShTu} + O(\varepsilon^{4/3}) \quad \text{and} \quad \left(r + \frac{\mu^2}{2\gamma\sigma^2}\right) - \text{ESR} = \frac{3}{4}\varepsilon\text{WeTu} + O(\varepsilon^{4/3}). \quad (3.3)$$

These two relations have the same meaning: the welfare effect of small transaction costs is proportional to trading volume times the spread. The constant of proportionality $3/4$ is universal, that is, independent of both investment opportunities (r, μ, σ) and preferences (γ) .

In the first formula, the welfare effect is measured by the liquidity premium, that is in terms of the risky asset. Likewise, trading volume is expressed as share turnover, which also focuses on the risky asset alone. By contrast, the second formula considers the decrease in the equivalent safe rate and wealth turnover, two quantities that treat both assets equally. In summary, if both welfare and volume are measured consistently with each other, the welfare effect approximately equals volume times the spread, up to the universal factor $3/4$.

Figure 4 plots the spread, share turnover, and the liquidity premium implied by the first equation in (3.3). As in Lo and Wang (2000), the spread and share turnover are capitalization-weighted averages of all securities in the CRSP monthly stocks database with share codes 10 and 11, and with nonzero bid, ask, volume and share outstanding. While turnover figures are available before 1992, separate bid and ask prices were not recorded until then, thereby preventing a reliable estimation of spreads for earlier periods.

Spreads steadily decline in the observation period, dropping by almost an order of magnitude after stock market decimalization of 2001. At the same time, trading volume substantially increases from a typical monthly turnover of 6% in the early 1990s to over 20% in the late 2000s. The implied liquidity premium also declines with spreads after decimalization, but less than the spread, in view of the increase in turnover. During the months of the financial crisis in late 2008, the implied liquidity premium rises sharply, not because of higher volumes, but because spreads widen substantially. Thus, although this implied liquidity premium is only a coarse estimate, it has advantages over other proxies, because it combines information on both prices and quantities, and is supported by a model.

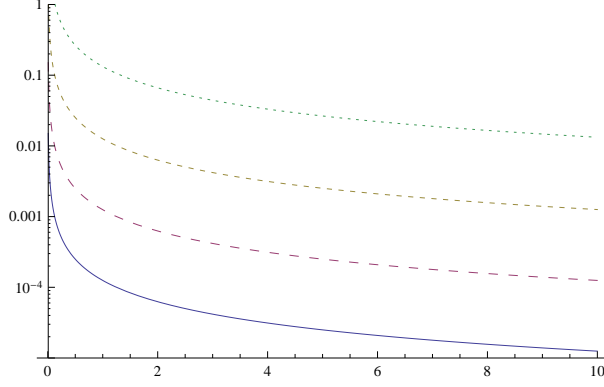


Figure 5: Upper bound on the difference between the long-run and finite-horizon equivalent safe rates (vertical axis), against the horizon (horizontal axis), for spread $\varepsilon = 0.01\%$ (solid), 0.1% (long dashed), 1% (short dashed), and 10% (dotted). Parameters are $\mu = 8\%$, $\sigma = 16\%$, $\gamma = 5$.

3.6 Finite Horizons

The trading boundaries in this paper are optimal for a long investment horizon, but are also approximately optimal for finite horizons. The following theorem, which complements the main result, makes this point precise:

Theorem 3.1. *Fix a time horizon $T > 0$. Then the finite-horizon equivalent safe rate of any strategy (ϕ^0, ϕ) satisfies the upper bound*

$$\frac{1}{T} \log E \left[(\Xi_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \leq r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2} + \frac{1}{T} \log(\phi_{0-}^0 + \phi_{0-} - S_0) + \pi_* \frac{\varepsilon}{T} + O(\varepsilon^{4/3}), \quad (3.4)$$

and the finite-horizon equivalent safe rate of our long-run optimal strategy (φ^0, φ) satisfies the lower bound

$$\frac{1}{T} \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \geq r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2} + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} - S_0) - \left(2\pi_* + \frac{\varphi_{0-} - S_0}{\varphi_{0-}^0 + \varphi_{0-} - S_0} \right) \frac{\varepsilon}{T} + O(\varepsilon^{4/3}). \quad (3.5)$$

In particular, for the same unlevered initial position ($\phi_{0-} = \varphi_{0-} \geq 0$, $\phi_{0-}^0 = \varphi_{0-}^0 \geq 0$), the equivalent safe rates of (ϕ^0, ϕ) and of the optimal policy (φ^0, φ) for horizon T differ by at most

$$\frac{1}{T} \left(\log E \left[(\Xi_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} - \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \right) \leq (3\pi_* + 1) \frac{\varepsilon}{T} + O(\varepsilon^{4/3}). \quad (3.6)$$

This result implies that the horizon, like consumption, only has a second order effect on portfolio choice with transaction costs, because the finite-horizon equivalent safe rate matches, at the order $\varepsilon^{2/3}$, the equivalent safe rate of the stationary long-run optimal policy. This result recovers, in particular, the first-order asymptotics for the finite-horizon value function obtained by Bichuch (2011, Theorem 4.1). In addition, Theorem 3.1 provides explicit estimates for the correction terms of order ε arising from liquidation costs. Indeed, $r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2}$ is the maximum rate achieved by trading optimally. The remaining terms arise due to the transient influence of the initial endowment, as well as the costs of the initial transaction, which takes place if the initial position lies outside the no-trade region, and of the final portfolio liquidation. These costs are of order ε/T because they are incurred

only once, and hence defrayed by a longer trading period. By contrast, portfolio rebalancing generates recurring costs, proportional to the horizon, and their impact on the equivalent safe rate does not decline as the horizon increases.

Even after accounting for all such costs in the worst-case scenario, the bound in (3.6) shows that their combined effect on the equivalent safe rate is lower than the spread ε , as soon as the horizon exceeds $3\pi_* + 1$, that is four years in the absence of leverage. Yet, this bound holds only up a term of order $\varepsilon^{4/3}$, so it is worth comparing it with the exact bounds in equations (C.17)-(C.18), from which (3.4) and (3.5) are obtained.

The exact bounds in Figure 5 show that, for typical parameter values, loss in equivalent safe rate of the long-run optimal strategy is lower than the spread ε even for horizons as short as 18 months, and quickly declines to become ten times smaller, for horizons close to ten years. In summary, the long-run approximation is a useful modeling device that makes the model tractable, and the resulting optimal policies are also nearly optimal even for horizons of a few years.

4 Heuristic Solution

This section contains an informal derivation of the main results. First, formal arguments of stochastic control are used to obtain the optimal policy, its welfare, and their asymptotic expansions.

4.1 Transaction costs market

For a trading strategy (φ_t^0, φ_t) , again write the number of shares $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ as the difference of the cumulated units purchased and sold, and denote by

$$X_t^0 = \varphi_t^0 S_t^0, \quad X_t = \varphi_t S_t,$$

the values of the safe and risky positions in terms of the ask price S_t . Then, the self-financing condition (2.1), and the dynamics of S_t^0 and S_t imply

$$\begin{aligned} dX_t^0 &= rX_t^0 dt - S_t d\varphi_t^\uparrow + (1 - \varepsilon)S_t d\varphi_t^\downarrow, \\ dX_t &= (\mu + r)X_t dt + \sigma X_t dW_t + S_t d\varphi_t^\uparrow - S_t d\varphi_t^\downarrow. \end{aligned}$$

Consider the maximization of expected power utility $U(x) = x^{1-\gamma}/(1-\gamma)$ from terminal wealth at time T , and denote by $V(t, x, y)$ its value function, which depends on time and the value of the safe and risky positions. Itô's formula yields:

$$\begin{aligned} dV(t, X_t^0, X_t) &= V_t dt + V_x dX_t^0 + V_y dX_t + \frac{1}{2} V_{yy} d\langle X, X \rangle_t \\ &= \left(V_t + rX_t^0 V_x + (\mu + r)X_t V_y + \frac{\sigma^2}{2} X_t^2 V_{yy} \right) dt \\ &\quad + S_t (V_y - V_x) d\varphi_t^\uparrow + S_t ((1 - \varepsilon)V_x - V_y) d\varphi_t^\downarrow + \sigma X_t V_y dW_t, \end{aligned}$$

where the arguments of the functions are omitted for brevity. Because $V(t, X_t^0, X_t)$ must be a supermartingale for any choice of the cumulative purchases and sales $\varphi_t^\uparrow, \varphi_t^\downarrow$ (which are increasing processes), it follows that $V_y - V_x \leq 0$ and $(1 - \varepsilon)V_x - V_y \leq 0$, that is

$$1 \leq \frac{V_x}{V_y} \leq \frac{1}{1 - \varepsilon}.$$

In the interior of this region, the drift of $V(t, X_t^0, X_t)$ cannot be positive, and must become zero for the optimal policy,

$$V_t + rX_t^0 V_x + (\mu + r)X_t V_y + \frac{\sigma^2}{2} X_t^2 V_{yy} = 0 \quad \text{if} \quad 1 < \frac{V_x}{V_y} < \frac{1}{1 - \varepsilon}. \quad (4.1)$$

To simplify further, note that the value function must be homogeneous with respect to wealth, and that — in the long run — it should grow exponentially with the horizon at a constant rate. These arguments lead to guess¹¹ that $V(t, X_t^0, X_t) = (X_t^0)^{1-\gamma} v(X_t/X_t^0) e^{-(1-\gamma)(r+\beta)t}$ for some β to be found. Setting $z = y/x$, the HJB equation reduces to

$$\frac{\sigma^2}{2} z^2 v''(z) + \mu z v'(z) - (1 - \gamma)\beta v(z) = 0 \quad \text{if} \quad 1 + z < \frac{(1 - \gamma)v(z)}{v'(z)} < \frac{1}{1 - \varepsilon} + z. \quad (4.2)$$

Assuming that the set $\{z : 1 + z \leq \frac{(1-\gamma)v(z)}{v'(z)} < \frac{1}{1-\varepsilon} + z\}$ coincides with some interval $l \leq z \leq u$ to be determined, and noting that at l the left inequality in (4.2) holds as equality, while at u the right inequality holds as equality, the following free boundary problem arises:

$$\frac{\sigma^2}{2} z^2 v''(z) + \mu z v'(z) - (1 - \gamma)\beta v(z) = 0 \quad \text{if } l < z < u, \quad (4.3)$$

$$(1 + l)v'(l) - (1 - \gamma)v(l) = 0, \quad (4.4)$$

$$(1/(1 - \varepsilon) + u)v'(u) - (1 - \gamma)v(u) = 0. \quad (4.5)$$

These conditions are not enough to identify the solution, because they can be matched for any choice of the trading boundaries l, u . The optimal boundaries are the ones that also satisfy the smooth-pasting conditions (cf. Dumas (1991)), formally obtained by differentiating (4.4) and (4.5) with respect to l and u , respectively:

$$(1 + l)v''(l) + \gamma v'(l) = 0, \quad (4.6)$$

$$(1/(1 - \varepsilon) + u)v''(u) + \gamma v'(u) = 0. \quad (4.7)$$

In addition to the reduced value function v , this system requires to solve for the excess equivalent safe rate β and the trading boundaries l and u . Substituting (4.6) and (4.4) into (4.3) yields (cf. Dumas and Luciano (1991))

$$-\frac{\sigma^2}{2}(1 - \gamma)\gamma \frac{l^2}{(1 + l)^2} v + \mu(1 - \gamma) \frac{l}{1 + l} v - (1 - \gamma)\beta v = 0.$$

Setting $\pi_- = l/(1 + l)$, and factoring out $(1 - \gamma)v$, it follows that

$$-\frac{\gamma\sigma^2}{2}\pi_-^2 + \mu\pi_- - \beta = 0.$$

Note that π_- is the risky weight when it is time to buy, and hence the risky position is valued at the ask price. The same argument for u shows that the other solution of the quadratic equation is $\pi_+ = u(1 - \varepsilon)/(1 + u(1 - \varepsilon))$, which is the risky weight when it is time to sell, and hence the risky position is valued at the bid price. Thus, the optimal policy is to buy when the “ask” fraction falls

¹¹This guess assumes that the cash position is strictly positive, $X_t^0 > 0$, which excludes leverage. With leverage, factoring out $(-X_t^0)^{1-\gamma}$ leads to analogous calculations.

below π_- , sell when the “bid” fraction rises above π_+ , and do nothing in between. Since π_- and π_+ solve the same quadratic equation, they are related to β via

$$\pi_{\pm} = \frac{\mu}{\gamma\sigma^2} \pm \frac{\sqrt{\mu^2 - 2\beta\gamma\sigma^2}}{\gamma\sigma^2}.$$

It is convenient to set $\beta = (\mu^2 - \lambda^2)/2\gamma\sigma^2$, because $\beta = \mu^2/2\gamma\sigma^2$ without transaction costs. We call λ the *gap*, since $\lambda = 0$ in a frictionless market, and, as λ increases, all variables diverge from their frictionless values. Put differently, to compensate for transaction costs, the investor would require another asset, with expected return λ and volatility σ , which trades without frictions and is uncorrelated with the risky asset.¹² With this notation, the buy and sell boundaries are just

$$\pi_{\pm} = \frac{\mu \pm \lambda}{\gamma\sigma^2}.$$

In other words, the buy and sell boundaries are symmetric around the classical frictionless solution $\mu/\gamma\sigma^2$. Since $l(\lambda), u(\lambda)$ are identified by π_{\pm} in terms of λ , it now remains to find λ . After deriving $l(\lambda)$ and $u(\lambda)$, the boundaries in the problem (4.3)-(4.5) are no longer free, but fixed. With the substitution

$$v(z) = e^{(1-\gamma)\int_0^{\log(z/l(\lambda))} w(y)dy}, \quad \text{i.e.,} \quad w(y) = \frac{l(\lambda)e^y v'(l(\lambda)e^y)}{(1-\gamma)v(l(\lambda)e^y)},$$

the boundary problem (4.3)-(4.5) reduces to a Riccati ODE

$$w'(x) + (1-\gamma)w(x)^2 + \left(\frac{2\mu}{\sigma^2} - 1\right)w(x) - \gamma\left(\frac{\mu - \lambda}{\gamma\sigma^2}\right)\left(\frac{\mu + \lambda}{\gamma\sigma^2}\right) = 0, \quad x \in [0, \log u(\lambda)/l(\lambda)], \quad (4.8)$$

$$w(0) = \frac{\mu - \lambda}{\gamma\sigma^2}, \quad (4.9)$$

$$w(\log(u(\lambda)/l(\lambda))) = \frac{\mu + \lambda}{\gamma\sigma^2}, \quad (4.10)$$

where

$$\frac{u(\lambda)}{l(\lambda)} = \frac{1}{(1-\varepsilon)} \frac{\pi_+(1-\pi_-)}{\pi_-(1-\pi_+)} = \frac{1}{(1-\varepsilon)} \frac{(\mu + \lambda)(\mu - \lambda - \gamma\sigma^2)}{(\mu - \lambda)(\mu + \lambda - \gamma\sigma^2)}. \quad (4.11)$$

For each λ , the initial value problem (4.8)-(4.9) has a solution $w(\lambda, \cdot)$, and the correct value of λ is identified by the second boundary condition (4.10).

4.2 Asymptotics

The equation (4.10) does not have an explicit solution, but it is possible to obtain an asymptotic expansion for small transaction costs ($\varepsilon \sim 0$) using the implicit function theorem. To this end, write the boundary condition (4.10) as $f(\lambda, \varepsilon) = 0$, where:

$$f(\lambda, \varepsilon) = w(\lambda, \log(u(\lambda)/l(\lambda))) - \frac{\mu + \lambda}{\gamma\sigma^2}.$$

Of course, $f(0, 0) = 0$ corresponds to the frictionless case. The implicit function theorem then suggests that around zero $\lambda(\varepsilon)$ follows the asymptotics $\lambda(\varepsilon) \sim -\varepsilon f_{\varepsilon}/f_{\lambda}$, but the difficulty is that

¹²Recall that in a market with two uncorrelated assets with returns μ_1 and μ_2 , both with volatility σ , the maximum Sharpe ratio is $(\mu_1^2 + \mu_2^2)/\sigma^2$. That is, squared Sharpe ratios add across orthogonal shocks.

$f_\lambda = 0$, because λ is not of order ε . Heuristic arguments (Shreve and Soner, 1994; Rogers, 2004) suggest that λ is of order $\varepsilon^{1/3}$. Thus, setting $\lambda = \delta^{1/3}$ and $\hat{f}(\delta, \varepsilon) = f(\delta^{1/3}, \varepsilon)$, and computing the derivatives of the explicit formula for $w(\lambda, x)$ (cf. Lemma B.1) shows that:

$$\hat{f}_\varepsilon(0, 0) = -\frac{\mu(\mu - \gamma\sigma^2)}{\gamma^2\sigma^4}, \quad \hat{f}_\delta(0, 0) = \frac{4}{3\mu^2\sigma^2 - 3\gamma\mu\sigma^4}.$$

As a result:

$$\delta(\varepsilon) \sim -\frac{f_\varepsilon}{f_\delta}\varepsilon = \frac{3\mu^2(\mu - \gamma\sigma^2)^2}{4\gamma^2\sigma^2}\varepsilon \quad \text{whence} \quad \lambda(\varepsilon) \sim \left(\frac{3\mu^2(\mu - \gamma\sigma^2)^2}{4\gamma^2\sigma^2}\right)^{1/3}\varepsilon^{1/3}.$$

The asymptotic expansions of all other quantities then follow by Taylor expansion.

5 Conclusion

In a tractable model of transaction costs with one safe and one risky asset and constant investment opportunities, we have computed explicitly the optimal trading policy, its welfare, liquidity premium, and trading volume, for an investor with constant relative risk aversion and a long horizon.

The trading boundaries are symmetric around the Merton proportion, if each boundary is computed with the corresponding trading price. Both the liquidity premium and trading volume are small in the unlevered regime, but become substantial in the presence of leverage. For a small bid-ask spread, the liquidity premium is approximately equal to share turnover times the spread, times the universal constant $3/4$.

Trading boundaries depend on investment opportunities only through the mean variance ratio. The equivalent safe rate, the liquidity premium, and trading volume also depend only on the mean variance ratio if measured in business time.

Appendix

A Derivation of the Shadow Market

The key to justify the above heuristic arguments is to reduce the portfolio choice problem with transaction costs to another portfolio choice problem, without transaction costs, with the bid and ask prices replaced by a single shadow price \tilde{S}_t evolving within the bid-ask spread, which coincides with either price at times of trading, and which yields the same optimal policy. In the case of logarithmic utility, this approach was applied successfully by Kallsen and Muhle-Karbe (2010) and Gerhold, Muhle-Karbe and Schachermayer (2011b,a).

Definition A.1. *A shadow price is a frictionless price process \tilde{S}_t , lying within the bid-ask spread $((1 - \varepsilon)S_t \leq \tilde{S}_t \leq S_t$ a.s.), such that there is an optimal strategy for \tilde{S}_t which is of finite variation, and entails buying only when the shadow price \tilde{S}_t equals the ask price S_t , and selling only when \tilde{S}_t equals the bid price $(1 - \varepsilon)S_t$.*

Once a candidate for such a shadow price is identified, long-run verification results for frictionless models (cf. Guasoni and Robertson (2012)) deliver the optimality of the guessed policy. Further, this method provides explicit upper and lower bounds on finite-horizon performance (cf. Lemma C.2 below), thereby allowing to check whether the long-run optimal strategy is approximately optimal for an horizon T . Put differently, it shows which horizons are long enough.

Finding a shadow price requires another heuristic argument. The idea is that \tilde{S}_t/S_t , the ratio between the shadow and the ask price, should only depend on the ratio of risky and safe weights, i.e., on the same state variable as the reduced value function in the transaction cost problem. Consequently, we look for a shadow price of the form¹³

$$\tilde{S}_t = \frac{S_t}{e^{Y_t}} g(e^{Y_t}),$$

where $e^{Y_t} = (X_t/X_t^0)/l$ is the ratio between the risky and safe positions at the ask price S_t , and centered at the buying boundary

$$l = \frac{\pi_-}{1 - \pi_-} = \frac{\mu - \lambda}{\gamma\sigma^2 - (\mu - \lambda)}.$$

Inside the no-trade region, the numbers of units φ_t^0 and φ_t remain constant so that $Y_t = \log(\varphi_t/l\varphi_t^0) + \log(S_t/S_t^0)$ follows Brownian motion with drift. Since Y_t must remain in $[0, \log(u/l)]$ by definition, it is reflected at the boundaries, that is,

$$dY_t = (\mu - \sigma^2/2)dt + \sigma dW_t + dL_t - dU_t,$$

for nondecreasing local time processes L_t, U_t that only increase on $\{Y_t = 0\}$ (resp. $\{Y_t = \log(u/l)\}$). The function $g : [1, u/l] \rightarrow [1, (1 - \varepsilon)u/l]$ is a C^2 -function satisfying the boundary conditions (cf. Gerhold, Muhle-Karbe and Schachermayer (2011b))

$$g(1) = 1, \quad g(u/l) = (1 - \varepsilon)u/l, \quad g'(1) = 1, \quad g'(u/l) = 1 - \varepsilon. \quad (\text{A.1})$$

The first two conditions ensure that \tilde{S}_t equals the ask price S_t (resp. the bid price $(1 - \varepsilon)S_t$) when Y_t sits at the buying boundary 0 (resp. at the selling boundary $\log(u/l)$). The boundary conditions for g' , and Itô's formula imply that \tilde{S}_t is an Itô process with dynamics

$$d\tilde{S}_t/\tilde{S}_t = (\tilde{\mu}(Y_t) + r)dt + \tilde{\sigma}(Y_t)dW_t,$$

where

$$\tilde{\mu}(y) = \frac{\mu g'(e^y)e^y + \frac{\sigma^2}{2}g''(e^y)e^{2y}}{g(e^y)}, \quad \text{and} \quad \tilde{\sigma}(y) = \frac{\sigma g'(e^y)e^y}{g(e^y)}.$$

To identify the function g , first derive the HJB equation for a generic g . Then, compare this equation to the one obtained in the previous section for the market with transaction costs. The function g is identified by the condition that the value functions of the two problems must be the same.

The wealth process corresponding to a policy $\tilde{\pi}_t$ in terms of the shadow price \tilde{S}_t is

$$d\tilde{X}_t = r\tilde{X}_t dt + \tilde{\pi}_t \tilde{\mu}(Y_t) \tilde{X}_t dt + \tilde{\pi}_t \tilde{\sigma}(Y_t) \tilde{X}_t dW_t.$$

The usual ansatz for the value function \tilde{V} in frictionless markets driven by a state variable Y_t , i.e. $\tilde{V} = \tilde{V}(t, \tilde{X}_t, Y_t)$, and Itô's formula yield

$$\begin{aligned} d\tilde{V}(t, \tilde{X}_t, Y_t) = & \left(\tilde{V}_t + r\tilde{X}_t \tilde{V}_x + \tilde{\mu} \tilde{\pi}_t \tilde{X}_t \tilde{V}_x + \frac{\tilde{\sigma}^2}{2} \tilde{\pi}_t^2 \tilde{X}_t^2 \tilde{V}_{xx} + \left(\mu - \frac{\sigma^2}{2} \right) \tilde{V}_y + \frac{\sigma^2}{2} \tilde{V}_{yy} + \sigma \tilde{\sigma} \tilde{\pi}_t \tilde{X}_t \tilde{V}_{xy} \right) dt \\ & + \tilde{V}_y (dL_t - dU_t) + (\tilde{\sigma} \tilde{\pi}_t \tilde{X}_t \tilde{V}_x + \sigma \tilde{V}_y) dW_t, \end{aligned}$$

¹³An equivalent guess is $\tilde{S}_t = S_t h(Y_t)$. With hindsight, the one in the text leads to simpler calculations, because S_t is a multiple of e^{Y_t} in the no-trade region.

where the arguments of the functions are omitted for brevity. Since \tilde{V} must be a supermartingale for any strategy, and a martingale for the optimal strategy, the HJB equation reads as

$$\sup_{\tilde{\pi}} \left(\tilde{V}_t + rx\tilde{V}_x + \tilde{\mu}\tilde{\pi}x\tilde{V}_x + \frac{\tilde{\sigma}^2}{2}\tilde{\pi}^2x^2\tilde{V}_{xx} + \left(\mu - \frac{\sigma^2}{2}\right)\tilde{V}_y + \frac{\sigma^2}{2}\tilde{V}_{yy} + \sigma\tilde{\sigma}\tilde{\pi}x\tilde{V}_{xy} \right) = 0,$$

with the Neumann boundary conditions

$$\tilde{V}_y(0) = \tilde{V}_y(\log(u/l)) = 0.$$

The homogeneity of the value function (i.e., $\tilde{V}(t, x, y) = x^{1-\gamma}\tilde{v}(t, y)$) leads to the first-order condition:

$$\tilde{\pi}_t = \frac{1}{\gamma} \left(\frac{\tilde{\mu}}{\tilde{\sigma}^2} + \frac{\sigma}{\tilde{\sigma}} \frac{\tilde{v}_y}{\tilde{v}} \right).$$

Plugging this equality back into the HJB equation yields the nonlinear equation

$$\tilde{v}_t + (1-\gamma)r\tilde{v} + \left(\mu - \frac{\sigma^2}{2}\right)\tilde{v}_y + \frac{\sigma^2}{2}\tilde{v}_{yy} + \frac{1-\gamma}{2\gamma} \left(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma \frac{\tilde{v}_y}{\tilde{v}} \right)^2 \tilde{v} = 0.$$

Now, the equivalent safe rate $\beta = (\mu^2 - \lambda^2)/2\gamma\sigma^2$ must be the same, both for the shadow market and for the transaction cost market in the previous section. Thus, setting

$$\tilde{v}(t, y) = e^{-(1-\gamma)(r+\beta)t} e^{(1-\gamma) \int_0^y \tilde{w}(z) dz},$$

which implies that $\tilde{v}_y/\tilde{v} = (1-\gamma)\tilde{w}$, the HJB equation reduces to the inhomogeneous Riccati ODE

$$\tilde{w}' + (1-\gamma)\tilde{w}^2 + \left(\frac{2\mu}{\sigma^2} - 1\right)\tilde{w} - \frac{2\beta}{\sigma^2} + \frac{1}{\gamma\sigma^2} \left(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1-\gamma)\tilde{w} \right)^2 = 0 \quad (\text{A.2})$$

with boundary conditions

$$\tilde{w}(0) = \tilde{w}(\log(u/l)) = 0. \quad (\text{A.3})$$

For \tilde{S}_t to be a shadow price, its value function

$$\tilde{V}_t = e^{-(1-\gamma)(r+\beta)t} \tilde{X}_t^{1-\gamma} e^{(1-\gamma) \int_0^y \tilde{w}(z) dz}$$

must coincide with the value function

$$V_t = e^{-(1-\gamma)(r+\beta)t} (X_t^0)^{1-\gamma} e^{(1-\gamma) \int_0^y w(z) dz}$$

for the transaction cost problem derived above. By definition, the safe position X_t^0 and the wealth \tilde{X}_t in terms of $\tilde{S}_t = S_t e^{-Y_t} g(e^{Y_t})$ are related via

$$\frac{\tilde{X}_t}{X_t^0} = \frac{\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t}{\varphi_t^0 S_t^0} = 1 + g(e^{Y_t})l.$$

Now, the condition $\tilde{V} = V$ implies that $0 = \log(1 + g(e^y)l) + \int_0^y (\tilde{w}(z) - w(z)) dz$, which in turn means that

$$\tilde{w}(y) = w(y) - \frac{g'(e^y)e^y l}{1 + g(e^y)l}. \quad (\text{A.4})$$

Plugging this relation into the ODE (A.2) for \tilde{w} , using the ODE (4.8) for w , and simplifying leads to

$$\left((1-\gamma)w(y) + \frac{\tilde{\mu}(y)}{\sigma\tilde{\sigma}(y)} - \frac{g'(e^y)e^y l}{1 + g(e^y)l} \right)^2 = 0.$$

Inserting the definitions of $\tilde{\mu}(y)$ and $\tilde{\sigma}(y)$, this relation is tantamount to the following ODE for g :¹⁴

$$\frac{g''(e^y)e^y}{g'(e^y)} - \frac{2g'(e^y)e^yl}{1+g(e^y)l} + \frac{2\mu}{\sigma^2} + 2(1-\gamma)w(y) = 0. \quad (\text{A.5})$$

Next, the substitution

$$k(y) = \frac{1+g(e^y)l}{g'(e^y)e^yl}, \quad \text{i.e.,} \quad g(e^y) = \left(1 + \frac{1}{l}\right) \exp\left(\int_0^y \frac{1}{k(z)} dz\right) - \frac{1}{l},$$

reduces this ODE to the inhomogeneous *linear* equation

$$k'(y) = k(y) \left(\frac{2\mu}{\sigma^2} - 1 + 2(1-\gamma)w(y) \right) - 1. \quad (\text{A.6})$$

Since $1/k(0) = w(0) - \tilde{w}(0)$ by (A.4), the boundary condition (A.3) for \tilde{w} and its counterpart (4.9) for w imply that $k(0) = \gamma\sigma^2/(\mu - \lambda)$. The solution to (A.6) then follows from the variation of constants formula. In each of the three different cases of Lemma B.1, plugging in the respective explicit formula for w (cf. Lemma B.1) and integrating leads to an explicit formula for k . In Case 2, we have (with constants a and b as in Lemma B.1)

$$k(y) = \cos^2 \left[\tan^{-1} \left(\frac{b}{a} \right) + ay \right] \left(-\frac{1}{a} \tan \left[\tan^{-1} \left(\frac{b}{a} \right) + ay \right] + \frac{b}{a^2} + \frac{(a^2 + b^2)\gamma\sigma^2}{a^2(\mu - \lambda)} \right),$$

The other two cases lead to analogous results (cf. Lemma B.4). Now the chain of substitutions is reversed starting from k , which is known explicitly up to the constant λ . First, set $\tilde{w}(y) = w(y) - 1/k(y)$; then $\tilde{w}(0) = 0$ by construction. To establish the other boundary condition $\tilde{w}(\log(u/l)) = 0$, it suffices to check that $1/k(\log(u/l)) = (\mu + \lambda)/(\gamma\sigma^2)$. To this end, insert the boundary conditions for w ,

$$\frac{\mu + \lambda}{\gamma\sigma^2} = w(\log(u/l)) = \frac{a \tan[\tan^{-1}(\frac{b}{a}) + a \log(\frac{u}{l})] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, \quad (\text{A.7})$$

$$\frac{\mu + \lambda}{\gamma\sigma^2} - \left(\frac{\mu + \lambda}{\gamma\sigma^2} \right)^2 = w'(\log(u/l)) = \frac{a^2}{(\gamma - 1) \cos^2[\tan^{-1}(\frac{b}{a}) + a \log(\frac{u}{l})]}, \quad (\text{A.8})$$

into the explicit formula for $1/k(y)$.¹⁵ Now, observe that the function

$$\begin{aligned} g(e^y) &= \left(1 + \frac{1}{l}\right) \exp\left(\int_0^y \frac{1}{k(z)} dz\right) - \frac{1}{l} \\ &= 1 + \frac{\gamma\sigma^2}{\mu - \lambda} \left(\frac{1}{1 + \frac{\mu - \lambda}{\gamma\sigma^2} \left(\frac{b}{a^2 + b^2} - \frac{a}{a^2 + b^2} \tan \left[\tan^{-1} \left(\frac{b}{a} \right) + ay \right] \right)} - 1 \right) \end{aligned}$$

satisfies $g(1) = 1$. Moreover, $g(u/l) = (1 - \varepsilon)u/l$, which again follows by inserting (A.8) into the explicit expression for g . Finally, these boundary conditions for g and those for k imply that $g'(1) = 1$ and $g'(u/l) = 1 - \varepsilon$, i.e., g satisfies the smooth pasting conditions (A.1) and, by construction, also the ODE (A.5).

¹⁴For logarithmic utility ($\gamma = 1$) this ODE reduces to the one in Gerhold, Muhle-Karbe and Schachermayer (2011b, Equation (3.5)).

¹⁵The first equalities in (A.7) (resp. (A.8)) follow from Equations (4.6) and (4.7) respectively. The second equalities follow from the explicit formula for w from Lemma B.1.

In summary, although the derivation and the formulas are more involved, the shadow price is determined as explicitly as in the case of logarithmic utility (Gerhold, Muhle-Karbe and Schachermayer, 2011b). That is, its dynamics are known explicitly in terms of λ , which is identified as the solution of a scalar equation and can be developed into a fractional power series in terms of $\varepsilon^{1/3}$.

With the heuristically derived candidate shadow price at hand, the proof of Theorem 2.2 is divided into three parts. The first part establishes that the explicit formulas for the reduced value function w , the auxiliary function k , and the function g parametrizing the shadow price are well-defined and indeed have the properties derived above. In the second part, these results are used to construct a shadow price, and to show that it satisfies the upper and lower finite-horizon bounds in Lemma C.2, which in turn imply long-run optimality. Finally, the third part contains the explicit calculation of the implied trading volume.

B Explicit formulas and their properties

The first step is to determine, for a given small $\lambda > 0$, an explicit expression for the solution w of the ODE (4.8), complemented by the initial condition (4.9).

Lemma B.1. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then for sufficiently small $\lambda > 0$, the function*

$$w(\lambda, x) = \begin{cases} \frac{a(\lambda) \tanh[\tanh^{-1}(b(\lambda)/a(\lambda)) - a(\lambda)x] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{if } \gamma \in (0, 1) \text{ and } \frac{\mu}{\gamma\sigma^2} < 1 \text{ or } \gamma > 1 \text{ and } \frac{\mu}{\gamma\sigma^2} > 1, \\ \frac{a(\lambda) \tan[\tan^{-1}(b(\lambda)/a(\lambda)) + a(\lambda)x] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{if } \gamma > 1 \text{ and } \frac{\mu}{\gamma\sigma^2} \in \left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{1}{\gamma}}, \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{1}{\gamma}}\right), \\ \frac{a(\lambda) \coth[\coth^{-1}(b(\lambda)/a(\lambda)) - a(\lambda)x] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{otherwise,} \end{cases}$$

with

$$a(\lambda) = \sqrt{\left|(\gamma - 1)\frac{\mu^2 - \lambda^2}{\gamma\sigma^4} - \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2\right|} \quad \text{and} \quad b(\lambda) = \frac{1}{2} - \frac{\mu}{\sigma^2} + (\gamma - 1)\frac{\mu - \lambda}{\gamma\sigma^2},$$

is a local solution of

$$w'(x) + (1 - \gamma)w^2(x) + \left(\frac{2\mu}{\sigma^2} - 1\right)w(x) - \frac{\mu^2 - \lambda^2}{\gamma\sigma^4} = 0, \quad w(0) = \frac{\mu - \lambda}{\gamma\sigma^2}. \quad (\text{B.1})$$

Moreover, $x \mapsto w(\lambda, x)$ is increasing (resp. decreasing) for $\mu/\gamma\sigma^2 \in (0, 1)$ (resp. $\mu/\gamma\sigma^2 > 1$).

Proof. The first part of the assertion is easily verified by taking derivatives. The second follows by inspection of the explicit formulas. \square

Next, establish that the crucial constant λ , which determines both the no-trade region and the equivalent safe rate, is well-defined.

Lemma B.2. *Let $0 < \mu/\gamma\sigma^2 \neq 1$ and $w(\lambda, \cdot)$ be defined as in Lemma B.1, and set*

$$l(\lambda) = \frac{\mu - \lambda}{\gamma\sigma^2 - (\mu - \lambda)}, \quad u(\lambda) = \frac{1}{(1 - \varepsilon)} \frac{\mu + \lambda}{\gamma\sigma^2 - (\mu + \lambda)}.$$

Then, for sufficiently small $\varepsilon > 0$, there exists a unique solution λ of

$$w\left(\lambda, \log\left(\frac{u(\lambda)}{l(\lambda)}\right)\right) - \frac{\mu + \lambda}{\gamma\sigma^2} = 0. \quad (\text{B.2})$$

As $\varepsilon \downarrow 0$, it has the asymptotics

$$\lambda = \gamma\sigma^2 \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{1/3} \varepsilon^{1/3} + \sigma^2 \left(\frac{(5-2\gamma)}{10} \frac{\mu}{\gamma\sigma^2} \left(1 - \frac{\mu}{\gamma\sigma^2} \right) - \frac{3}{20} \right) \varepsilon + O(\varepsilon^{4/3}).$$

Proof. The explicit expression for w in Lemma B.1 implies that $w(\lambda, x)$ in Lemma B.1 is analytic in both variables at $(0, 0)$. By the initial condition in (B.1), its power series has the form

$$w(\lambda, x) = \frac{\mu - \lambda}{\gamma\sigma^2} + \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} W_{ij} x^i \lambda^j,$$

where expressions for the coefficients W_{ij} are computed by expanding the explicit expression for w . Hence, the left-hand side of the boundary condition (B.2) is an analytic function of ε and λ . Its power series expansion shows that the coefficients of $\varepsilon^0 \lambda^j$ vanish for $j = 0, 1, 2$, so that the condition (B.2) reduces to

$$\lambda^3 \sum_{i \geq 0} A_i \lambda^i = \varepsilon \sum_{i, j \geq 0} B_{ij} \varepsilon^i \lambda^j \quad (\text{B.3})$$

with (computable) coefficients A_i and B_{ij} . This equation has to be solved for λ . Since

$$A_0 = \frac{4}{3\mu\sigma^2(\gamma\sigma^2 - \mu)} \quad \text{and} \quad B_{00} = \frac{\mu(\gamma\sigma^2 - \mu)}{\gamma^2\sigma^4}$$

are non-zero, divide the equation (B.3) by $\sum_{i \geq 0} A_i \lambda^i$, and take the third root, obtaining that, for some C_{ij} ,

$$\lambda = \varepsilon^{1/3} \sum_{i, j \geq 0} C_{ij} \varepsilon^i \lambda^j = \varepsilon^{1/3} \sum_{i, j \geq 0} C_{ij} (\varepsilon^{1/3})^{3i} \lambda^j.$$

The right-hand side is an analytic function of λ and $\varepsilon^{1/3}$, so that the implicit function theorem (Gunning and Rossi, 2009, Theorem I.B.4) yields a unique solution λ (for ε sufficiently small), which is an analytic function of $\varepsilon^{1/3}$. Its power series coefficients can be computed at any order. \square

Henceforth, consider small transaction costs $\varepsilon > 0$, and let λ denote the constant in Lemma B.2. Moreover, set $w(x) = w(\lambda, x)$, $a = a(\lambda)$, $b = b(\lambda)$, and $u = u(\lambda)$, $l = l(\lambda)$. By inspection, it follows that the function w satisfies the following smooth pasting conditions:

Lemma B.3. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then, in all three cases,*

$$w'(0) = \frac{\mu - \lambda}{\gamma\sigma^2} - \left(\frac{\mu - \lambda}{\gamma\sigma^2} \right)^2, \quad w' \left(\log \left(\frac{u}{l} \right) \right) = \frac{\mu + \lambda}{\gamma\sigma^2} - \left(\frac{\mu + \lambda}{\gamma\sigma^2} \right)^2.$$

The next lemma states the properties of the function k .

Lemma B.4. *Let $0 < \mu/\gamma\sigma^2 \neq 1$ and define*

$$k(y) = \begin{cases} \cosh^2 \left[\tanh^{-1} \left(\frac{b}{a} \right) - ay \right] \left(\frac{1}{a} \tanh \left[\tanh^{-1} \left(\frac{b}{a} \right) - ay \right] - \frac{b}{a^2} + \frac{(a^2 - b^2)\gamma\sigma^2}{a^2(\mu - \lambda)} \right), \\ \cos^2 \left[\tan^{-1} \left(\frac{b}{a} \right) + ay \right] \left(-\frac{1}{a} \tan \left[\tan^{-1} \left(\frac{b}{a} \right) + ay \right] + \frac{b}{a^2} + \frac{(a^2 + b^2)\gamma\sigma^2}{a^2(\mu - \lambda)} \right), \\ \sinh^2 \left[\coth^{-1} \left(\frac{b}{a} \right) - ay \right] \left(-\frac{1}{a} \coth \left[\coth^{-1} \left(\frac{b}{a} \right) - ay \right] + \frac{b}{a^2} + \frac{(b^2 - a^2)\gamma\sigma^2}{a^2(\mu - \lambda)} \right), \end{cases}$$

with the same cases as in Lemma B.1. Then k satisfies the linear ODE

$$k'(y) = k(y) \left(\frac{2\mu}{\sigma^2} - 1 + 2(1 - \gamma)w(y) \right) - 1, \quad 0 \leq y \leq \log\left(\frac{u}{l}\right),$$

with boundary conditions

$$k(0) = \frac{\gamma\sigma^2}{\mu - \lambda}, \quad k\left(\log\left(\frac{u}{l}\right)\right) = \frac{\gamma\sigma^2}{\mu + \lambda}.$$

Moreover, k is strictly decreasing (resp. increasing) and, in particular, strictly positive on $[0, \log(u/l)]$ for $\mu/\gamma\sigma^2 \in (0, 1)$ (resp. on $[\log(u/l), 0]$ for $\mu/\gamma\sigma^2 > 1$).

Proof. That k satisfies the ODE follows by insertion. The identities $\cos^2[\tan^{-1}(x)] = 1/(1 + x^2)$ as well as $\cosh^2[\tanh^{-1}(x)] = 1/(1 - x^2)$ and $\sinh^2[\coth^{-1}(x)] = 1/(x^2 - 1)$ yield the boundary condition at zero, whereas the boundary condition at $\log(u/l)$ follows by inserting

$$w\left(\log\left(\frac{u}{l}\right)\right) = \frac{\mu + \lambda}{\gamma\sigma^2} \quad \text{and} \quad w'\left(\log\left(\frac{u}{l}\right)\right) = \frac{\mu + \lambda}{\gamma\sigma^2} - \left(\frac{\mu + \lambda}{\gamma\sigma^2}\right)^2.$$

Finally, the ODE and a comparison argument yield that k is strictly decreasing (resp. increasing for $\mu/\gamma\sigma^2 > 1$). \square

Lemma B.5. Let $0 < \mu/\gamma\sigma^2 \neq 1$ and define

$$g(e^y) := \frac{\gamma\sigma^2}{\mu - \lambda} \exp\left(\int_0^y \frac{1}{k(z)} dz\right) - \left(\frac{\gamma\sigma^2}{\mu - \lambda} - 1\right),$$

for $0 \leq y \leq \log(u/l)$ if $\mu/\gamma\sigma^2 \in (0, 1)$ resp. for $\log(u/l) \leq y \leq 0$ if $\mu/\gamma\sigma^2 > 1$. Then

$$g(e^y) = \begin{cases} 1 + \frac{\gamma\sigma^2}{\mu - \lambda} \left(\left(1 + \frac{\mu - \lambda}{\gamma\sigma^2} \left(\frac{b}{b^2 - a^2} - \frac{a}{b^2 - a^2} \tanh\left[\tanh^{-1}\left(\frac{b}{a}\right) - ay\right] \right) \right)^{-1} - 1 \right), \\ 1 + \frac{\gamma\sigma^2}{\mu - \lambda} \left(\left(1 + \frac{\mu - \lambda}{\gamma\sigma^2} \left(\frac{b}{b^2 + a^2} - \frac{a}{a^2 + b^2} \tanh\left[\tanh^{-1}\left(\frac{b}{a}\right) + ay\right] \right) \right)^{-1} - 1 \right), \\ 1 + \frac{\gamma\sigma^2}{\mu - \lambda} \left(\left(1 + \frac{\mu - \lambda}{\gamma\sigma^2} \left(\frac{b}{b^2 - a^2} - \frac{a}{b^2 - a^2} \coth\left[\coth^{-1}\left(\frac{b}{a}\right) - ay\right] \right) \right)^{-1} - 1 \right), \end{cases}$$

with the same cases as in Lemma B.1, and g satisfies the boundary and smooth pasting conditions

$$g(1) = 1, \quad g(u/l) = (1 - \varepsilon)u/l, \quad g'(1) = 1, \quad g'(u/l) = 1 - \varepsilon.$$

Moreover, $g' > 0$, hence g maps $[1, u/l]$ (resp. $[u/l, 1]$) onto $[1, (1 - \varepsilon)u/l]$ (resp. $[(1 - \varepsilon)u/l, 1]$) if $\mu/\gamma\sigma^2 > 1$). Finally, g solves

$$\frac{g''(e^y)e^y}{g'(e^y)} - \frac{2g'(e^y)e^yl}{1 + g(e^y)l} + \frac{2\mu}{\sigma^2} + 2(1 - \gamma)w(y) = 0. \quad (\text{B.4})$$

Proof. The explicit representation follows by elementary integration. Evidently, $g(1) = 1$. Moreover, $g(u/l) = (1 - \varepsilon)u/l$ follows by inserting $(\mu + \lambda)/\gamma\sigma^2 = w(\log(u/l))$ once again. Next, $g(1) = 1$, $g(u/l) = (1 - \varepsilon)u/l$, $k(0) = \gamma\sigma^2/(\mu - \lambda)$, and $k(\log(u/l)) = \gamma\sigma^2/(\mu + \lambda)$, and

$$g'(e^y) = \frac{e^{-y}}{k(y)} \left(g(e^y) + \frac{\gamma\sigma^2}{\mu - \lambda} - 1 \right) \quad (\text{B.5})$$

imply $g'(1) = 1$ and $g'(u/l) = 1 - \varepsilon$. Hence g satisfies the smooth pasting conditions. Furthermore, (B.5) and a comparison argument yield that $g' > 0$ because $k > 0$. Finally, computing the derivatives shows that g indeed satisfies the ODE (B.4). \square

C The shadow price and verification

To construct the shadow price as in Gerhold, Muhle-Karbe and Schachermayer (2011*b,a*), for $y \in [0, \log(u/l)]$ (resp. $[\log(u/l), 0]$ if $\mu/\gamma\sigma^2 > 1$), let Y_t be Brownian motion with drift, reflected at 0 and $\log(u/l)$, that is, the continuous, adapted process with values in $[0, \log(u/l)]$ (resp. in $[\log(u/l), 0]$ if $\mu/\gamma\sigma^2 > 1$), such that

$$dY_t = (\mu - \sigma^2/2)dt + \sigma dW_t + dL_t - dU_t, \quad Y_0 = y, \quad (\text{C.1})$$

for nondecreasing (resp. nonincreasing if $\mu/\gamma\sigma^2 > 1$) adapted processes L_t and U_t increasing (resp. decreasing if $\mu/\gamma\sigma^2 > 1$) only on the sets $\{Y_t = 0\}$ and $\{Y_t = \log(u/l)\}$, respectively.

Lemma C.1. *Define*

$$y = \begin{cases} 0, & \text{if } l\xi^0 S_0^0 \geq \xi S_0, \\ \log(u/l), & \text{if } u\xi^0 S_0^0 \leq \xi S_0, \\ \log[(\xi S_0/\xi^0 S_0^0)/l], & \text{otherwise,} \end{cases} \quad (\text{C.2})$$

and let Y_t be reflected Brownian motion with drift as in (C.1), started at $Y_0 = y$. Then $\tilde{S}_t = S_t e^{-Y_t} g(e^{Y_t})$, with g as in Lemma B.5, is a positive Itô process with dynamics

$$d\tilde{S}_t/\tilde{S}_t = (\tilde{\mu}(Y_t) + r)dt + \tilde{\sigma}(Y_t)dW_t, \quad \tilde{S}_0 = S_0 e^{-y} g(e^y),$$

for

$$\tilde{\mu}(y) = \frac{\mu g'(e^y)e^y + \frac{\sigma^2}{2}g''(e^y)e^{2y}}{g(e^y)}, \quad \tilde{\sigma}(y) = \frac{\sigma g'(e^y)e^y}{g(e^y)},$$

and \tilde{S}_t takes values in the bid-ask spread $[(1 - \varepsilon)S_t, S_t]$.

Note that the first (resp. second) case in (C.2) occurs if the initial ratio $\xi S_0/\xi^0 S_0^0$ lies below the buying boundary l (resp. above the selling boundary l for $\mu/\gamma\sigma^2 > 1$) or above the selling boundary u (resp. below the buying boundary u for $\mu/\gamma\sigma^2 > 1$). Then, there is a jump from the initial position $(\varphi_{0-}^0, \varphi_{0-}) = (\xi^0, \xi)$, which moves the ratio to the nearest boundary of the interval $[l, u]$ (cf. Lemma C.3 below). Since this initial trade involves the purchase (resp. sale) of shares, the initial value of \tilde{S}_t is chosen to match the initial ask (resp. bid) price.

Proof of Lemma C.1. The first part of the assertion follows from the smooth pasting conditions for g and Itô's formula. As for the second part, since $g''(1) \leq 0$, a comparison argument yields that the derivative $(g'(y)y - g(y))/y^2$ of $g(y)/y$ is non-positive. Hence $g(1)/1 = 1$ and $g(u/l)/(u/l) = 1 - \varepsilon$ yield that $\tilde{S}_t = S_t g(e^{Y_t})e^{-Y_t}$ is indeed $[(1 - \varepsilon)S_t, S_t]$ -valued. \square

The long-run optimal portfolio in the frictionless “shadow market” with price process \tilde{S}_t can now be determined by adapting the argument in Guasoni and Robertson (2012).

Lemma C.2. *The function $\tilde{w}(y) = w(y) - g'(e^y)e^y l/(1 + g(e^y)l)$ solves the system (A.2)-(A.3), that is:*

$$\frac{\sigma^2}{2}\tilde{w}' + (1 - \gamma)\frac{\sigma^2}{2}\tilde{w}^2 + \left(\mu - \frac{\sigma^2}{2}\right)\tilde{w} + \frac{1}{2\gamma}\left(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1 - \gamma)\tilde{w}\right)^2 = \beta, \quad (\text{C.3})$$

with boundary conditions $\tilde{w}(0) = \tilde{w}(\log(u/l)) = 0$ and $\beta = (\mu^2 - \lambda^2)/2\gamma\sigma^2$. Setting $\tilde{q}(y) = \int_0^y \tilde{w}(z)dz$, the shadow payoff \tilde{X}_T corresponding to the policy $\tilde{\pi} = \frac{1}{\gamma}\left(\frac{\tilde{\mu}}{\tilde{\sigma}^2} + (1 - \gamma)\frac{\sigma}{\tilde{\sigma}}\tilde{w}\right)$ (in terms of

\tilde{S}_t) and the shadow discount factor $\tilde{M}_T = e^{-rT} \mathcal{E}(-\int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW_t)_T$ satisfy the following bounds:

$$E \left[\tilde{X}_T^{1-\gamma} \right] = \tilde{X}_0^{1-\gamma} e^{(1-\gamma)(r+\beta)T} \hat{E} \left[e^{(1-\gamma)(\tilde{q}(Y_0) - \tilde{q}(Y_T))} \right], \quad (\text{C.4})$$

$$E \left[\tilde{M}_T^{1-\frac{1}{\gamma}} \right]^\gamma = e^{(1-\gamma)(r+\beta)T} \hat{E} \left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(Y_0) - \tilde{q}(Y_T))} \right]^\gamma. \quad (\text{C.5})$$

Here, $\hat{E}[\cdot]$ denotes the expectation with respect to the myopic probability \hat{P} , defined by

$$\frac{d\hat{P}}{dP} = \exp \left(\int_0^T \left(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}\tilde{\pi} \right) dW_t - \frac{1}{2} \int_0^T \left(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}\tilde{\pi} \right)^2 dt \right).$$

Proof. First note that $\tilde{\mu}, \tilde{\sigma}, \tilde{\pi}, \tilde{w}$ are functions of Y_t , but the argument is omitted throughout to ease notation. Next, it is readily verified by insertion that \tilde{w} satisfies the boundary value problem. Now, to prove (C.4), notice that the shadow wealth process \tilde{X}_t satisfies:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} e^{\int_0^T (r + \tilde{\mu}\tilde{\pi} - \frac{\tilde{\sigma}^2}{2}\tilde{\pi}^2) dt + (1-\gamma) \int_0^T \tilde{\sigma}\tilde{\pi} dW_t}.$$

Hence:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{\int_0^T ((1-\gamma)(r + \tilde{\mu}\tilde{\pi} - \frac{\tilde{\sigma}^2}{2}\tilde{\pi}^2) + \frac{1}{2}(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}\tilde{\pi})^2) dt} e^{\int_0^T ((1-\gamma)\tilde{\sigma}\tilde{\pi} - (-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}\tilde{\pi})) dW_t}.$$

Substituting $\tilde{\pi} = \frac{1}{\gamma}(\frac{\tilde{\mu}}{\tilde{\sigma}^2} + (1-\gamma)\frac{\sigma}{\tilde{\sigma}}\tilde{w})$, the second integrand simplifies to $-(1-\gamma)\sigma\tilde{w}$. Similarly, the first integrand first reduces to $(1-\gamma)r + \frac{1}{2}\frac{\tilde{\mu}^2}{\tilde{\sigma}^2} + \gamma\frac{\tilde{\sigma}^2}{2}\tilde{\pi}^2 - \gamma\tilde{\mu}\tilde{\pi}$, and then to the expression $(1-\gamma)r + (1-\gamma)^2\frac{\sigma^2}{2}\tilde{w}^2 + \frac{1-\gamma}{2\gamma}(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1-\gamma)\tilde{w})^2$. In summary:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{(1-\gamma) \int_0^T (r + (1-\gamma)\frac{\sigma^2}{2}\tilde{w}^2 + \frac{1}{2\gamma}(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1-\gamma)\tilde{w})^2) dt - (1-\gamma) \int_0^T \sigma\tilde{w} dW_t}. \quad (\text{C.6})$$

Now, Itô's formula and the boundary conditions $\tilde{w}(0) = \tilde{w}(\log(u/l)) = 0$ imply that:

$$\begin{aligned} \tilde{q}(Y_T) - \tilde{q}(Y_0) &= \int_0^T \tilde{w}(Y_t) dY_t + \frac{1}{2} \int_0^T \tilde{w}'(Y_t) d\langle Y, Y \rangle_t \\ &= \int_0^T \left(\left(\mu - \frac{\sigma^2}{2} \right) \tilde{w} + \frac{\sigma^2}{2} \tilde{w}' \right) dt + \int_0^T \sigma \tilde{w} dW_t. \end{aligned}$$

Using this identity to replace the term $\int_0^T \sigma \tilde{w} dW_t$ in (C.6) yields

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{(1-\gamma) \int_0^T (r + \frac{\sigma^2}{2}\tilde{w}' + (1-\gamma)\frac{\sigma^2}{2}\tilde{w}^2 + (\mu - \frac{\sigma^2}{2})\tilde{w} + \frac{1}{2\gamma}(\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1-\gamma)\tilde{w})^2) dt} e^{-(1-\gamma)(\tilde{q}(Y_T) - \tilde{q}(Y_0))}.$$

Since \tilde{w} satisfies (C.3), the first bound follows:

$$E \left[\tilde{X}_T^{1-\gamma} \right] = \tilde{X}_0^{1-\gamma} E \left[\frac{d\hat{P}}{dP} e^{(1-\gamma)(r+\beta)T - (1-\gamma)(\tilde{q}(Y_T) - \tilde{q}(Y_0))} \right] = \tilde{X}_0^{1-\gamma} e^{(1-\gamma)(r+\beta)T} \hat{E} \left[e^{(\gamma-1)(\tilde{q}(Y_T) - \tilde{q}(Y_0))} \right].$$

The argument for the second bound is similar. The shadow discount factor $\tilde{M}_T = e^{-rT} \mathcal{E}(-\int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW)_T$ and the myopic probability \hat{P} satisfy

$$\tilde{M}_T^{1-\frac{1}{\gamma}} = e^{\frac{1-\gamma}{\gamma} \int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW_t + \frac{1-\gamma}{\gamma} \int_0^T (r + \frac{\tilde{\mu}^2}{2\tilde{\sigma}^2}) dt}, \quad \frac{d\hat{P}}{dP} = e^{\frac{1-\gamma}{\gamma} \int_0^T (\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma\tilde{w}) dW_t - \frac{(1-\gamma)^2}{2\gamma^2} \int_0^T (\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma\tilde{w})^2 dt}.$$

Hence:

$$\tilde{M}_T^{1-\frac{1}{\gamma}} = \frac{d\hat{P}}{dP} e^{-\frac{1-\gamma}{\gamma} \int_0^T \sigma \tilde{w} dW_t + \frac{1-\gamma}{\gamma} \int_0^T r + \frac{1}{2} (\frac{\tilde{\mu}^2}{\tilde{\sigma}^2} + \frac{1-\gamma}{\gamma} (\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma \tilde{w})^2) dt}.$$

Since $\frac{\tilde{\mu}^2}{\tilde{\sigma}^2} + \frac{1-\gamma}{\gamma} (\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma \tilde{w})^2 = (1-\gamma)\sigma^2 \tilde{w}^2 + \frac{1}{\gamma} (\frac{\tilde{\mu}}{\tilde{\sigma}} + \sigma(1-\gamma)\tilde{w})^2$, substituting again the equality

$$\int_0^T \sigma \tilde{w} dW_t = \tilde{q}(Y_T) - \tilde{q}(Y_0) - \int_0^T \left(\left(\mu - \frac{\sigma^2}{2} \right) \tilde{w} + \frac{\sigma^2}{2} \tilde{w}' \right) dt$$

and the HJB equation (C.3) yields

$$\tilde{M}_T^{1-\frac{1}{\gamma}} = \frac{d\hat{P}}{dP} e^{\frac{1-\gamma}{\gamma} (r+\beta)T - \frac{1-\gamma}{\gamma} (\tilde{q}(Y_T) - \tilde{q}(Y_0))}.$$

The second bound then follows by taking the expectation, and raising it to power of γ . \square

With the finite horizon bounds at hand, it is now straightforward to establish that the policy $\tilde{\pi}(Y_t)$ is indeed long-run optimal in the frictionless market with price \tilde{S}_t .

Lemma C.3. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then, the policy*

$$\tilde{\pi}(Y_t) = \frac{1}{\gamma} \left(\frac{\tilde{\mu}(Y_t)}{\tilde{\sigma}^2(Y_t)} + (1-\gamma) \frac{\sigma}{\tilde{\sigma}(Y_t)} \tilde{w}(Y_t) \right) = \frac{g(e^{Y_t})l}{1 + g(e^{Y_t})l} \quad (\text{C.7})$$

is long-run optimal with equivalent safe rate $r + \beta$ in the frictionless market with price process \tilde{S}_t . The corresponding wealth process (in terms of \tilde{S}_t), and the numbers of safe and risky units are given by

$$\begin{aligned} \tilde{X}_t &= (\xi^0 S_0^0 + \xi \tilde{S}_0) \mathcal{E} \left(\int_0^t (r + \tilde{\pi}(Y_t) \tilde{\mu}(Y_t)) dt + \int_0^t \tilde{\pi}(Y_t) \tilde{\sigma}(Y_t) dW_t \right), \\ \varphi_{0-} &= \xi, \quad \varphi_t = \tilde{\pi}(Y_t) \tilde{X}_t / \tilde{S}_t \quad \text{for } t \geq 0, \\ \varphi_{0-}^0 &= \xi^0, \quad \varphi_t^0 = (1 - \tilde{\pi}(Y_t)) \tilde{X}_t / S_t^0 \quad \text{for } t \geq 0. \end{aligned}$$

Proof. The second representation for $\tilde{\pi}(Y_t)$ follows by inserting the definitions of $\tilde{\mu}(y), \tilde{\sigma}(y)$ from Lemma C.1, the ODE (B.4) for $g(y)$, and the identity $1/k(y) = g'(e^y)e^y l / (1 + g(e^y)l)$, which is a direct consequence of the definition of g . The formulas for the corresponding wealth process and the numbers of safe and risky units follow from the standard frictionless definitions. Now let \tilde{M}_t be the shadow discount factor from Lemma C.2. Then, standard duality arguments for power utility (cf. Lemma 5 in Guasoni and Robertson (2012)) imply that the shadow payoff \tilde{X}_t^ϕ corresponding to *any* admissible strategy ϕ_t satisfies the inequality

$$E \left[(\tilde{X}_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \leq E \left[\tilde{M}_T^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{1-\gamma}}. \quad (\text{C.8})$$

This inequality in turn yields the following upper bound, valid for any admissible strategy ϕ_t in the frictionless market with shadow price \tilde{S}_t :

$$\liminf_{T \rightarrow \infty} \frac{1}{(1-\gamma)T} \log E \left[(\tilde{X}_T^\phi)^{1-\gamma} \right] \leq \liminf_{T \rightarrow \infty} \frac{\gamma}{(1-\gamma)T} \log E \left[\tilde{M}_T^{\frac{\gamma-1}{\gamma}} \right]. \quad (\text{C.9})$$

Since the function \tilde{q} is bounded on the compact support of Y_t , the second bound in Lemma C.2 implies that the right-hand side equals $r + \beta$. Likewise, the first bound in the same lemma implies that the shadow payoff \tilde{X}_t (corresponding to the policy φ_t) attains this upper bound, concluding the proof. \square

The next Lemma establishes that \tilde{S}_t is a shadow price. Here the argument is similar to the one for logarithmic utility (Gerhold, Muhle-Karbe and Schachermayer, 2011a).

Lemma C.4. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then, the number of shares $\varphi_t = \tilde{\pi}(Y_t)\tilde{X}_t/\tilde{S}_t$ in the portfolio $\tilde{\pi}(Y_t)$ in Lemma C.3 has the dynamics*

$$\frac{d\varphi_t}{\varphi_t} = \left(1 - \frac{\mu - \lambda}{\gamma\sigma^2}\right) dL_t - \left(1 - \frac{\mu + \lambda}{\gamma\sigma^2}\right) dU_t. \quad (\text{C.10})$$

Thus, for $\mu/\gamma\sigma^2 \in (0, 1)$, φ_t increases only when $Y_t = 0$, that is, when \tilde{S}_t equals the ask price, and decreases only when $Y_t = \log(u/l)$, that is, when \tilde{S}_t equals the bid price.

Proof. Itô's formula applied to (C.7) yields

$$\begin{aligned} \frac{d\tilde{\pi}(Y_t)}{\tilde{\pi}(Y_t)} &= \frac{\mu e^{Y_t}(1 + g(e^{Y_t})l)g'(e^{Y_t}) - \sigma^2 l e^{2Y_t}g'(e^{Y_t})^2 + \frac{1}{2}\sigma^2 e^{2Y_t}(1 + g(e^{Y_t})l)g''(e^{Y_t})}{g(e^{Y_t})(1 + g(e^{Y_t})l)^2} dt \\ &\quad + \frac{\sigma g'(e^{Y_t})e^{Y_t}}{g(e^{Y_t})(1 + g(e^{Y_t})l)} dW_t + \frac{g'(e^{Y_t})e^{Y_t}}{g(e^{Y_t})(1 + g(e^{Y_t})l)} (dL_t - dU_t). \end{aligned}$$

Integrating $\varphi_t = \tilde{\pi}(Y_t)\tilde{X}_t/\tilde{S}_t$ by parts twice, using the dynamics of $\tilde{\pi}(Y_t)$, \tilde{X}_t , and \tilde{S}_t , and simplifying, it follows that

$$\frac{d\varphi_t}{\varphi_t} = \left(\frac{g'(e^{Y_t})e^{Y_t}}{g(e^{Y_t})(1 + g(e^{Y_t})l)} \right) d(L_t - U_t).$$

Since L_t and U_t only increase (resp. decrease when $\mu/\gamma\sigma^2 > 1$) on $\{Y_t = 0\}$ and $\{Y_t = \log(u/l)\}$, respectively, the assertion now follows from the boundary conditions for g and g' . \square

The optimal growth rate for any frictionless price within the bid-ask spread must be greater or equal than in the original market with bid-ask process $((1 - \varepsilon)S_t, S_t)$, because the investor trades at more favorable prices. For a *shadow price*, there is an optimal strategy that only entails buying (resp. selling) stocks when \tilde{S} coincides with the ask- resp. bid price. Hence, this strategy yields the same payoff when executed at bid-ask prices, and thus is also optimal in the original model with transaction costs. The corresponding equivalent safe rate must also be the same, since the difference due to the liquidation costs vanishes as the horizon grows in (2.2):

Proposition C.5. *Suppose \tilde{S}_t is a shadow price for a bid-ask process $((1 - \varepsilon)S_t, S_t)$, with long-run optimal portfolio (φ_t^0, φ_t) as in Definition A.1, and the corresponding policy $\tilde{\pi}_t = \varphi_t \tilde{S}_t / (\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t)$ is bounded by $C(1 - \varepsilon)/\varepsilon$ for some constant $C \in (0, 1)$. Then, the portfolio (φ_t^0, φ_t) is also long-run optimal for $((1 - \varepsilon)S_t, S_t)$, with the same equivalent safe rate as for the frictionless price \tilde{S} .*

Proof. As φ_t only increases (resp. decreases) when $\tilde{S}_t = S_t$ (resp. $\tilde{S}_t = (1 - \varepsilon)S_t$), (φ_t^0, φ_t) is also self-financing for the bid-ask process $((1 - \varepsilon)S_t, S_t)$. In addition, since $S_t \geq \tilde{S}_t \geq (1 - \varepsilon)S_t$,

$$\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t \geq \varphi_t^0 S_t^0 + \varphi_t^+(1 - \varepsilon)S_t - \varphi_t^- S_t \geq \left(1 - \frac{\varepsilon}{1 - \varepsilon} |\tilde{\pi}_t|\right) (\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t). \quad (\text{C.11})$$

Thus, if (φ_t^0, φ_t) is admissible for \tilde{S}_t , it is also admissible for $((1 - \varepsilon)S_t, S_t)$, because $|\tilde{\pi}_t|\varepsilon/(1 - \varepsilon) \in (0, 1)$. Moreover, since $|\tilde{\pi}_t|\varepsilon/(1 - \varepsilon)$ is even bounded away from 1 by assumption, (C.11) also yields

$$\begin{aligned} \liminf_{T \rightarrow \infty} \frac{1}{(1 - \gamma)T} \log E \left[(\varphi_T^0 S_T^0 + \varphi_T^+(1 - \varepsilon)S_T - \varphi_T^- S_T)^{1 - \gamma} \right] \\ = \liminf_{T \rightarrow \infty} \frac{1}{(1 - \gamma)T} \log E \left[(\varphi_T^0 S_T^0 + \varphi_T \tilde{S}_T)^{1 - \gamma} \right], \end{aligned} \quad (\text{C.12})$$

that is, (φ_t^0, φ_t) has the same growth rate, either with \tilde{S}_t or with $[(1 - \varepsilon)S_t, S_t]$.

For any admissible strategy (ψ_t^0, ψ_t) for the bid-ask spread $[(1 - \varepsilon)S_t, S_t]$, set $\tilde{\psi}_t^0 = \psi_{0-}^0 - \int_0^t \tilde{S}_s/S_s^0 d\psi_s$. Then, $(\tilde{\psi}_t^0, \psi_t)$ is a self-financing trading strategy for \tilde{S}_t with $\tilde{\psi}_t^0 \geq \psi_t^0$. Together with $\tilde{S}_t \in [(1 - \varepsilon)S_t, S_t]$, the long-run optimality of (φ_t^0, φ_t) for \tilde{S}_t , and (C.12), it follows that:

$$\begin{aligned} & \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\psi_T^0 S_T^0 + \psi_T^+(1 - \varepsilon)S_T - \psi_T^- S_T)^{1-\gamma}] \\ & \leq \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\tilde{\psi}_T^0 S_T^0 + \psi_T \tilde{S}_T)^{1-\gamma}] \\ & \leq \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\varphi_T^0 S_T^0 + \varphi_T \tilde{S}_T)^{1-\gamma}] \\ & = \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\varphi_T^0 S_T^0 + \varphi_T^+(1 - \varepsilon)S_T - \varphi_T^- S_T)^{1-\gamma}]. \end{aligned}$$

Hence (φ_t^0, φ_t) is also long-run optimal for $((1 - \varepsilon)S_t, S_t)$. \square

The main result now follows by putting together the above statements.

Theorem C.6. *For $\varepsilon > 0$ small, and $0 < \mu/\gamma\sigma^2 \neq 1$, the process \tilde{S}_t in Lemma C.1 is a shadow price. A long-run optimal policy — both for the frictionless market with price \tilde{S}_t and in the market with bid-ask prices $(1 - \varepsilon)S_t, S_t$ — is to keep the risky weight $\tilde{\pi}_t$ (in terms of \tilde{S}_t) in the no-trade region*

$$[\pi_-, \pi_+] = \left[\frac{g(1)l}{1 + g(1)l}, \frac{g(u/l)l}{1 + g(u/l)l} \right] = \left[\frac{\mu - \lambda}{\gamma\sigma^2}, \frac{\mu + \lambda}{\gamma\sigma^2} \right].$$

As $\varepsilon \downarrow 0$, its boundaries have the asymptotics

$$\pi_{\pm} = \frac{\mu}{\gamma\sigma^2} \pm \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{1/3} \varepsilon^{1/3} \pm \left(\frac{(5 - 2\gamma)}{10\gamma} \frac{\mu}{\gamma\sigma^2} \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 - \frac{3}{20\gamma} \right) \varepsilon + O(\varepsilon^{4/3}).$$

The corresponding equivalent safe rate is:

$$r + \beta = r + \frac{\mu^2 - \lambda^2}{\gamma\sigma^2} = r + \frac{\mu^2}{2\gamma\sigma^2} - \frac{\gamma\sigma^2}{2} \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{2/3} \varepsilon^{2/3} + O(\varepsilon^{4/3}).$$

If $\mu/\gamma\sigma^2 = 1$, then $\tilde{S}_t = S_t$ is a shadow price, and it is optimal to invest all wealth in the risky asset at time $t = 0$, never to trade afterwards. In this case, the equivalent safe rate is given by the frictionless value $r + \beta = r + \mu^2/2\gamma\sigma^2$.

Proof. First let $0 < \mu/\gamma\sigma^2 \neq 1$. Optimality of the strategy (φ_t^0, φ_t) associated to $\tilde{\pi}(Y_t)$ for \tilde{S}_t has been shown in Lemma C.3. The second representation for π_{\pm} follows from the boundary conditions for g and the definition of u in Lemma B.2, while the asymptotic expansions are an immediate consequence of the fractional power series for λ (cf. Lemma B.2) and Taylor expansion.

Next, Lemma C.4 shows that \tilde{S}_t is a shadow price process in the sense of Definition A.1. In view of the asymptotic expansions for π_{\pm} , Proposition C.5 shows that, for small transaction costs ε , the same policy is also optimal, with the same equivalent safe rate, in the original market with bid-ask prices $(1 - \varepsilon)S_t, S_t$.

Consider now the degenerate case $\mu/\gamma\sigma^2 = 1$. Then the optimal strategy in the frictionless model $\tilde{S}_t = S_t$ transfers all wealth to the risky asset at time $t = 0$, never to trade afterwards

($\varphi_t^0 = 0$ and $\varphi_t = \xi + \xi^0 S_0^0/S_0$ for all $t \geq 0$). Hence it is of finite variation and the number of shares never decreases, and increases only at time $t = 0$, where the shadow price coincides with the ask price. Thus, $\tilde{S}_t = S_t$ is a shadow price. For small ε , the remaining assertions then follow from Proposition C.5 as above. \square

Next, the proof of Theorem 3.1, which establishes asymptotic finite-horizons bounds. In fact, the proof yields exact bounds in terms of λ , from which the expansions in the theorem are obtained.

Proof of Theorem 3.1. Let (ϕ^0, ϕ) be any admissible strategy starting from the initial position $(\varphi_{0-}^0, \varphi_{0-})$. Then as in the proof of Proposition C.5, we have $\Xi_T^\phi \leq \tilde{X}_T^\phi$ for the corresponding shadow payoff, that is, the terminal value of the wealth process $\tilde{X}_t^\phi = \phi_0^0 + \phi_0 \tilde{S}_0 + \int_0^t \phi_s d\tilde{S}_s$ corresponding to trading ϕ in the frictionless market with price process \tilde{S}_t . Hence (Guasoni and Robertson, 2012, Lemma 5) and the second bound in Lemma C.2 imply

$$\frac{1}{(1-\gamma)T} \log E \left[(\Xi_T^\phi)^{1-\gamma} \right] \leq r + \beta + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} S_0) + \frac{\gamma}{(1-\gamma)T} \log \hat{E} \left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(Y_0)-\tilde{q}(Y_T))} \right]. \quad (\text{C.13})$$

For the strategy (φ^0, φ) from Lemma C.4, we have $\Xi_T^\varphi \geq (1 - \frac{\varepsilon}{1-\varepsilon} \frac{\mu+\lambda}{\gamma\sigma^2}) \tilde{X}_T^\varphi$ by the proof of Proposition C.5. Hence the first bound in Lemma C.2 yields

$$\begin{aligned} \frac{1}{(1-\gamma)T} \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right] &\geq r + \beta + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} \tilde{S}_0) + \frac{1}{(1-\gamma)T} \log \hat{E} \left[e^{(1-\gamma)(\tilde{q}(Y_0)-\tilde{q}(Y_T))} \right] \\ &\quad + \frac{1}{T} \log \left(1 - \frac{\varepsilon}{1-\varepsilon} \frac{\mu+\lambda}{\gamma\sigma^2} \right). \end{aligned} \quad (\text{C.14})$$

To determine explicit estimates for these bounds, we first analyze the sign of $\tilde{w}(y)$ and hence the monotonicity of $\tilde{q}(y) = \int_0^y \tilde{w}(z) dz$. Whenever $\tilde{w}(y) = 0$, the ODE (C.3) yields

$$\sigma^2 \gamma \tilde{w}'(y) = \frac{\mu^2 - \lambda^2}{\sigma^2} - \frac{\tilde{\mu}(y)^2}{\tilde{\sigma}(y)^2}.$$

Plugging in the definitions of $\tilde{\mu}(y)$ and $\tilde{\sigma}(y)$ (cf. Lemma C.1) and using the ODE (B.4) for g , this equality reduces to

$$\sigma^2 \gamma \tilde{w}'(y) = \frac{\mu^2 - \lambda^2}{\sigma^2} - \left(\frac{\sigma g'(e^y) e^y l}{1 + g(e^y) l} \right)^2.$$

Now note that $\lambda = O(\varepsilon^{1/3})$ and $\log(u/l) = O(\varepsilon^{1/3})$ for $\varepsilon \downarrow 0$. Hence a Taylor expansion and the boundary conditions for g show

$$\sigma^2 \gamma \tilde{w}'(y) = \frac{\mu^2}{\sigma^2} \left(1 - \frac{1}{\gamma^2} \right) + O(\varepsilon^{1/3}).$$

Consequently — for a sufficiently small spread ε — the derivative of \tilde{w} is always positive for $\tilde{w}(y) = 0$ if $\gamma > 1$, resp. negative when $\gamma \in (0, 1)$. In view of $\tilde{w}(0) = 0$, a comparison argument then shows that $\tilde{w}(y) \geq 0$ for $\gamma > 1$ resp. $\tilde{w} \leq 0$ for $\gamma \in (0, 1)$. Hence the function $\tilde{q}(y)$ is increasing in the first case and decreasing in the second.

Now, consider Case 2 of Lemma B.1; the calculations for the other two cases follow along the same lines with minor modifications. Then $\gamma > 1$ and \tilde{q} is positive and increasing. Hence,

$$\frac{\gamma}{(1-\gamma)T} \log \hat{E} \left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(Y_0)-\tilde{q}(Y_T))} \right] \leq \frac{1}{T} \int_0^{\log(u/l)} \tilde{w}(y) dy \quad (\text{C.15})$$

and likewise

$$\frac{1}{(1-\gamma)T} \log \hat{E} \left[e^{(1-\gamma)(\tilde{q}(Y_0) - \tilde{q}(Y_T))} \right] \geq -\frac{1}{T} \int_0^{\log(u/l)} \tilde{w}(y) dy. \quad (\text{C.16})$$

Now, inserting $\tilde{w}(y) = w(y) - g'(e^y)e^y l / (1 + g(e^y)l)$ and integrating leads to

$$\int_0^{\log(u/l)} \tilde{w}(y) dy = \int_0^{\log(u/l)} w(y) dy - \log \left(\frac{\mu - \lambda - \gamma \sigma^2}{\mu + \lambda - \gamma \sigma^2} \right), \quad (\text{C.17})$$

due to the boundary conditions for g and the definition of u, l . By elementary integration of the explicit formula in Lemma B.1 and using the boundary conditions from Lemma B.3 for the evaluation of the result at 0 resp. $\log(u/l)$, the integral of w can also be computed in closed form:

$$\int_0^{\log(u/l)} w(y) dy = \frac{\frac{\mu}{\sigma^2} - \frac{1}{2}}{\gamma - 1} \log \left(\frac{1}{1 - \varepsilon} \frac{(\mu + \lambda)(\mu - \lambda - \gamma \sigma^2)}{(\mu - \lambda)(\mu + \lambda - \gamma \sigma^2)} \right) + \frac{1}{2(\gamma - 1)} \log \left(\frac{(\mu + \lambda)(\mu + \lambda - \gamma \sigma^2)}{(\mu - \lambda)(\mu - \lambda - \gamma \sigma^2)} \right). \quad (\text{C.18})$$

As $\varepsilon \downarrow 0$, a Taylor expansion and the power series for λ then yield

$$\int_0^{\log(u/l)} \tilde{w}(y) dy = \frac{\mu}{\gamma \sigma^2} \varepsilon + O(\varepsilon^{4/3}).$$

Likewise,

$$\log \left(1 - \frac{\varepsilon}{1 - \varepsilon} \frac{\mu - \lambda}{\gamma \sigma^2} \right) = -\frac{\mu}{\gamma \sigma^2} \varepsilon + O(\varepsilon^{4/3}),$$

as well as

$$\log(\varphi_{0-}^0 + \varphi_{0-} \tilde{S}_0) \geq \log(\varphi_{0-}^0 + \varphi_{0-} S_0) - \frac{\varphi_{0-} S_0}{\varphi_{0-}^0 + \varphi_{0-} S_0} \varepsilon + O(\varepsilon^2),$$

and the claimed bounds follow from (C.13) and (C.15) resp. (C.14) and (C.16). \square

D Trading volume

As above, let $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ denote the number of risky units at time t , written as the difference of the cumulated numbers of shares bought resp. sold until t . *Relative share turnover*, defined as the measure $d\|\varphi\|_t/|\varphi_t| = d\varphi_t^\uparrow/|\varphi_t| + d\varphi_t^\downarrow/|\varphi_t|$, is a scale-invariant indicator of trading volume (Lo and Wang, 2000). The *long-term average share turnover* is defined as

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{d\|\varphi\|_t}{|\varphi_t|}.$$

Similarly, *relative wealth turnover* $(1 - \varepsilon)S_t d\varphi_t^\downarrow / (\varphi_t^0 S_t^0 + \varphi_t(1 - \varepsilon)S_t) + S_t d\varphi_t^\uparrow / (\varphi_t^0 S_t^0 + \varphi_t S_t)$ is defined as the amount of wealth transacted divided by current wealth, where both quantities are evaluated in terms of the bid price $(1 - \varepsilon)S_t$ when selling shares resp. in terms of the ask price S_t when purchasing them. As above, the *long-term average wealth turnover* is then defined as

$$\lim_{T \rightarrow \infty} \frac{1}{T} \left(\int_0^T \frac{(1 - \varepsilon)S_t d\varphi_t^\downarrow}{\varphi_t^0 S_t^0 + \varphi_t(1 - \varepsilon)S_t} + \int_0^T \frac{S_t d\varphi_t^\uparrow}{\varphi_t^0 S_t^0 + \varphi_t S_t} \right).$$

Both of these limits admit explicit formulas in terms of the gap, which yield asymptotic expansions for $\varepsilon \downarrow 0$. The analysis starts with a preparatory result (cf. Janeček and Shreve (2004, Remark 4) for the case of driftless Brownian motion).

Lemma D.1. Let Y_t be a diffusion on an interval $[l, u]$, $0 < l < u$, reflected at the boundaries, i.e.

$$dY_t = b(Y_t)dt + a(Y_t)^{1/2}dW_t + dL_t - dU_t, \quad (\text{D.1})$$

where the mappings $a(y) > 0$ and $b(y)$ are both continuous, and the continuous, nondecreasing local time processes L_t and U_t satisfy $L_0 = U_0 = 0$ and only increase on $\{L_t = l\}$ and $\{U_t = u\}$, respectively. Denoting by $\nu(y)$ the invariant density of Y_t , the following almost sure limits hold:

$$\lim_{T \rightarrow \infty} \frac{L_T}{T} = \frac{a(l)\nu(l)}{2} \quad \lim_{T \rightarrow \infty} \frac{U_T}{T} = \frac{a(u)\nu(u)}{2} \quad (\text{D.2})$$

Proof. For $f \in C^2([l, u])$, write $\mathcal{L}f(y) := b(y)f'(y) + a(y)f''(y)/2$. Then, by Itô's formula:

$$\frac{f(Y_T) - f(Y_0)}{T} = \frac{1}{T} \int_0^T \mathcal{L}f(Y_t)dt + \frac{1}{T} \int_0^T f'(Y_t)a(Y_t)^{1/2}dW_t + f'(l)\frac{L_T}{T} - f'(u)\frac{U_T}{T}.$$

Now, take f such that $f'(l) = 1$ and $f'(u) = 0$, and pass to the limit $T \rightarrow \infty$. The left-hand side vanishes because f is bounded; the stochastic integral also vanishes by the Dambis-Dubins-Schwarz theorem, the law of the iterated logarithm, and the boundedness of f' . Thus, the ergodic theorem (Borodin and Salminen, 2002, II.36 and II.37) implies that

$$\lim_{T \rightarrow \infty} \frac{L_T}{T} = - \int_l^u \mathcal{L}f(y)\nu(y)dy.$$

Now, the self-adjoint representation (Revuz and Yor, 1999, VII.3.12) $\mathcal{L}f = (af'\nu)'/2\nu$ yields:

$$\lim_{T \rightarrow \infty} \frac{L_T}{T} = - \frac{1}{2} \int_l^u (af'\nu)'(y)dy = \frac{a(l)\nu(l)f'(l)}{2} - \frac{a(u)\nu(u)f'(u)}{2} = \frac{a(l)\nu(l)}{2}.$$

The other limit follows from the same argument, using f such that $f'(l) = 0$ and $f'(u) = 1$. \square

Lemma D.2. Let $0 < \mu/\gamma\sigma^2 \neq 1$ and, as in (C.1), let

$$Y_t = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t + L_t - U_t$$

be Brownian motion with drift, reflected at 0 and $\log(u/l)$. Then if $\mu \neq \sigma^2/2$, the following almost sure limits hold:

$$\lim_{T \rightarrow \infty} \frac{L_T}{T} = \frac{\sigma^2}{2} \left(\frac{\frac{2\mu}{\sigma^2} - 1}{(u/l)^{\frac{2\mu}{\sigma^2} - 1} - 1} \right) \quad \text{and} \quad \lim_{T \rightarrow \infty} \frac{U_T}{T} = \frac{\sigma^2}{2} \left(\frac{1 - \frac{2\mu}{\sigma^2}}{(u/l)^{1 - \frac{2\mu}{\sigma^2}} - 1} \right).$$

If $\mu = \sigma^2/2$, then $\lim_{T \rightarrow \infty} L_T/T = \lim_{T \rightarrow \infty} U_T/T = \sigma^2/(2\log(u/l))$ a.s.

Proof. First let $\mu \neq \sigma^2/2$. Moreover, suppose that $\mu/\gamma\sigma^2 \in (0, 1)$. Then the scale function and the speed measure of the diffusion Y_t are

$$s(x) = \int_0^x \exp\left(-2 \int_0^\xi \frac{\mu - \frac{\sigma^2}{2}}{\sigma^2} d\zeta\right) d\zeta = \frac{1}{1 - \frac{2\mu}{\sigma^2}} e^{(1 - \frac{2\mu}{\sigma^2})x},$$

$$m(dx) = 1_{[0, \log(u/l)]}(x) \frac{2dx}{s'(x)\sigma^2} = 1_{[0, \log(u/l)]}(x) \frac{2}{\sigma^2} e^{(\frac{2\mu}{\sigma^2} - 1)x} dx.$$

The invariant distribution of Y_t is the normalized speed measure

$$\nu(dx) = \frac{m(dx)}{m([0, \log(u/l)])} = 1_{[0, \log(u/l)]}(x) \frac{\frac{2\mu}{\sigma^2} - 1}{(u/l)^{\frac{2\mu}{\sigma^2} - 1} - 1} e^{(\frac{2\mu}{\sigma^2} - 1)x} dx.$$

For $\mu/\gamma\sigma^2 > 1$, the endpoints 0 and $\log(u/l)$ exchange their roles, and the result is the same, up to replacing $[0, \log(u/l)]$ with $[\log(u/l), 0]$ and multiplying the formula by -1 . Then, the claim follows from Lemma D.1. In the case $\mu = \sigma^2/2$ of driftless Brownian motion, Y_t has uniform stationary distribution on $[0, \log(u/l)]$ (resp. on $[\log(u/l), 0]$ if $\mu/\gamma\sigma^2 > 1$), and the claim again follows by Lemma D.1. \square

Lemma D.2 and the formula for φ from Lemma C.4 yield the long-term average trading volumes. The asymptotic expansions then follow from the power series for λ (cf. Lemma B.2).

Corollary D.3. *If $\mu/\gamma\sigma^2 \neq 1$, the long-term average share turnover is*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{d\|\varphi\|_t}{|\varphi_t|} = \left(1 - \frac{\mu - \lambda}{\gamma\sigma^2}\right) \lim_{T \rightarrow \infty} \frac{L_T}{T} + \left(1 - \frac{\mu + \lambda}{\gamma\sigma^2}\right) \lim_{T \rightarrow \infty} \frac{U_T}{T},$$

and the long-term average wealth turnover is

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \left(\int_0^T \frac{(1 - \varepsilon) S_t d\varphi_t^\downarrow}{\varphi_t^0 S_t^0 + \varphi_t(1 - \varepsilon) S_t} + \int_0^T \frac{S_t d\varphi_t^\uparrow}{\varphi_t^0 S_t^0 + \varphi_t S_t} \right) \\ = \frac{\mu - \lambda}{\gamma\sigma^2} \left(1 - \frac{\mu - \lambda}{\gamma\sigma^2}\right) \lim_{T \rightarrow \infty} \frac{L_T}{T} + \frac{\mu + \lambda}{\gamma\sigma^2} \left(1 - \frac{\mu + \lambda}{\gamma\sigma^2}\right) \lim_{T \rightarrow \infty} \frac{U_T}{T}, \end{aligned}$$

If $\mu/\gamma\sigma^2 = 1$, the long-term average share and wealth turnover both vanish.

References

- Ané, T. and Geman, H. (2000), ‘Order flow, transaction clock, and normality of asset returns’, *J. Finance* **55**(5), 2259–2284.
- Bichuch, M. (2011), ‘Asymptotic analysis for optimal investment in finite time with transaction costs’, *Siam J. Financial Math.* (To appear).
- Borodin, A. N. and Salminen, P. (2002), *Handbook of Brownian motion—facts and formulae*, Probability and its Applications, second edn, Birkhäuser Verlag, Basel.
- Constantinides, G. (1986), ‘Capital market equilibrium with transaction costs’, *J. Polit. Economy* **94**(4), 842–862.
- Dai, M. and Yi, F. (2009), ‘Finite-horizon optimal investment with transaction costs: a parabolic double obstacle problem’, *J. Differential Equations* **246**(4), 1445–1469.
- Davis, M. H. A. and Norman, A. R. (1990), ‘Portfolio selection with transaction costs’, *Math. Oper. Res.* **15**(4), 676–713.
- Dumas, B. (1991), ‘Super contact and related optimality conditions’, *J. Econom. Dynam. Control* **15**(4), 675–685.
- Dumas, B. and Luciano, E. (1991), ‘An exact solution to a dynamic portfolio choice problem under transactions costs’, *J. Finance* **46**(2), 577–595.
- Dybvig, P., Rogers, L. and Back, K. (1999), ‘Portfolio turnpikes’, *Rev. Finan. Stud.* **12**(1), 165–195.
- Gallant, A., Rossi, P. and Tauchen, G. (1992), ‘Stock prices and volume’, *Rev. Finan. Stud.* **5**(2), 199–242.
- Gerhold, S., Muhle-Karbe, J. and Schachermayer, W. (2011a), ‘Asymptotics and duality for the Davis and Norman problem’, *Stochastics* (To appear).
- Gerhold, S., Muhle-Karbe, J. and Schachermayer, W. (2011b), ‘The dual optimizer for the growth-optimal portfolio under transaction costs’, *Finance Stoch.* (To appear).

- Grossman, S. and Vila, J. (1992), ‘Optimal dynamic trading with leverage constraints’, *J. Finan. Quant. Anal.* **27**(2), 151–168.
- Grossman, S. and Zhou, Z. (1993), ‘Optimal investment strategies for controlling drawdowns’, *Math. Finance* **3**(3), 241–276.
- Guasoni, P. and Robertson, S. (2012), ‘Portfolios and risk premia for the long run’, *Ann. Appl. Probab.* **22**(1), 239–284.
- Gunning, R. C. and Rossi, H. (2009), *Analytic functions of several complex variables*, AMS Chelsea Publishing, Providence, RI.
- Janeček, K. and Shreve, S. E. (2004), ‘Asymptotic analysis for optimal investment and consumption with transaction costs’, *Finance Stoch.* **8**(2), 181–206.
- Kallsen, J. and Muhle-Karbe, J. (2010), ‘On using shadow prices in portfolio optimization with transaction costs’, *Ann. Appl. Probab.* **20**(4), 1341–1358.
- Liu, H. and Loewenstein, M. (2002), ‘Optimal portfolio selection with transaction costs and finite horizons’, *Rev. Finan. Stud.* **15**(3), 805–835.
- Lo, A. and Wang, J. (2000), ‘Trading volume: Definitions, data analysis, and implications of portfolio theory’, *Rev. Finan. Stud.* **13**(2), 257–300.
- Luttmer, E. (1996), ‘Asset pricing in economies with frictions’, *Econometrica* **64**(6), 1439–1467.
- Lynch, A. and Tan, S. (2011), ‘Explaining the magnitude of liquidity premia: The roles of return predictability, wealth shocks, and state-dependent transaction costs’, *J. Finance* **66**(4), 1329–1368.
- Magill, M. J. P. and Constantinides, G. M. (1976), ‘Portfolio selection with transactions costs’, *J. Econom. Theory* **13**(2), 245–263.
- Markowitz, H. (1952), ‘Portfolio Selection’, *J. Finance* **7**(1), 77–91.
- Merton, R. (1969), ‘Lifetime portfolio selection under uncertainty: The continuous-time case’, *Rev. Econ. Statist.* **51**(3), 247–257.
- Merton, R. C. (1971), ‘Optimum consumption and portfolio rules in a continuous-time model’, *J. Econom. Theory* **3**(4), 373–413.
- Revuz, D. and Yor, M. (1999), *Continuous martingales and Brownian motion*, third edn, Springer, Berlin.
- Rogers, L. C. G. (2004), Why is the effect of proportional transaction costs $O(\delta^{2/3})$?, in ‘Mathematics of finance’, Vol. 351 of *Contemp. Math.*, Amer. Math. Soc., Providence, RI, pp. 303–308.
- Shreve, S. E. and Soner, H. M. (1994), ‘Optimal investment and consumption with transaction costs’, *Ann. Appl. Probab.* **4**(3), 609–692.
- Taksar, M., Klass, M. J. and Assaf, D. (1988), ‘A diffusion model for optimal portfolio selection in the presence of brokerage fees’, *Math. Oper. Res.* **13**(2), 277–294.